



TUGAS AKHIR – SS141501

**PEMODELAN ANGKA PUTUS SEKOLAH USIA SMA
DI INDONESIA DENGAN PENDEKATAN REGRESI
NONPARAMETRIK SPLINE**

**HANIDA MURYANINGGAR
NRP 1313 100 090**

**Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR – SS141501

**PEMODELAN ANGKA PUTUS SEKOLAH USIA SMA
DI INDONESIA DENGAN PENDEKATAN REGRESI
NONPARAMETRIK SPLINE**

**HANIDA MURYANINGGAR
NRP 1313 100 090**

**Dosen Pembimbing
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT – SS141501

**MODELING OF HIGH SCHOOL DROPOUT
RATES IN INDONESIA USING NONPARAMETRIC
REGRESSION APPROACH SPLINE**

**HANIDA MURYANINGGAR
NRP 1313 100 090**

**Supervisor
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**

LEMBAR PENGESAHAN

PEMODELAN ANGKA PUTUS SEKOLAH USIA SMA DI INDONESIA DENGAN PENDEKATAN REGRESI NONPARAMETRIK SPLINE

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

Hanida Muryaninggar
NRP. 1313 100 090

Disetujui oleh Pembimbing:
Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
NIP. 19650603 198903 1 003



Mengetahui,
Kepala Departemen



Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JULI 2017

PEMODELAN ANGKA PUTUS SEKOLAH USIA SMA DI INDONESIA DENGAN PENDEKATAN REGRESI NONPARAMETRIK SPLINE

Nama Mahasiswa : Hanida Muryaninggar
NRP : 1313 100 090
Departemen : Statistika
**Dosen Pembimbing : Prof. Dr. I Drs. Nyoman
Budiantara, M. Si.**

Abstrak

Pendidikan merupakan hal yang sangat penting untuk meningkatkan kualitas sumber daya manusia. Salah satu program yang dilakukan pemerintah untuk meningkatkan taraf pendidikan di Indonesia adalah Program Indonesia Pintar (PIP) melalui usia wajib belajar 12 tahun. Pada kenyataannya, angka putus sekolah di Indonesia masih tinggi terutama pada jenjang pendidikan SMA. Angka putus sekolah pada usia SMA di Indonesia pada tahun 2016 mencapai 4,98 persen. Angka ini cukup tinggi bila dibandingkan dengan kelompok umur lainnya. Tingginya angka putus sekolah usia SMA di Indonesia disebabkan oleh banyak faktor. Oleh karena itu, dilakukan sebuah penelitian berupa pemodelan untuk mengetahui faktor-faktor yang mempengaruhi angka putus sekolah usia SMA di Indonesia. Pada penelitian ini, ada empat faktor yang diduga mempengaruhi angka putus sekolah usia SMA di Indonesia. Metode yang digunakan untuk memodelkan angka putus sekolah usia SMA dan faktor yang diduga mempengaruhinya adalah regresi nonparametrik spline. Metode ini dipilih dikarenakan Pola hubungan antara angka putus sekolah usia SMA di Indonesia dengan faktor-faktor yang diduga mempengaruhi tidak membentuk pola tertentu. Hasil penelitian ini menunjukkan bahwa model regresi nonparametrik spline terbaik adalah menggunakan tiga titik knot dengan 2 variabel yang signifikan yaitu persentase guru SM/Sederajat terhadap murid dan rata-rata lama sekolah penduduk usia 15 tahun ke atas. Nilai R^2 yang diperoleh sebesar 78,74%.

Kata kunci :Angka Putus Sekolah Usia SMA, Regresi Nonparametrik Spline, Titik Knot

MODELING OF HIGH SCHOOL DROPOUT RATES IN INDONESIA USING NONPARAMETRIC REGRESSION APPROACH SPLINE

Student's Name : Hanida Muryaninggar
NRP : 1313100090
Department : Statistika
**Supervisor : Prof. Dr. Drs. I Nyoman
Budiantara, M.Si.**

Abstract

Education is very important to improve the quality of human resources. One of the programs undertaken by the government to improve the level of education in Indonesia is the Program of Indonesia Smart (PIP) through 12 years compulsory education. But in reality, the dropout rate in Indonesia is still high especially in high school education level. The high school dropout rates in Indonesia in 2016 reaches 4.98 percent. This figure is quite high when compared with other age groups. The high school dropout rate in Indonesia is caused by many factors. Therefore, the researcher conducts a research in the form of modeling to determine the factors that influence the high school dropout rates in Indonesia. In this research, there are four factors that allegedly affect of high school dropout rates in Indonesia. The method used to model the high school drop-out rate and the factors that allegedly influence it is non-parametric Spline regression. Methods are selected because the Pattern of the relationship between high school dropout rates in Indonesia with factors that allegedly influence does not form a particular pattern. The results of this study indicate that the best non-parametric Spline regression model is using three of knots with 2 significant variables that is the percentage of high school teachers toward the students and the average length of school of people aged 15 years and over. The value of R^2 obtained is equal to 78,74%.

Keywords: High School Dropout Rate, Nonparametric Spline Regression, Knot Point

(This page intentionally left blank)

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan yang Maha Esa atas berkat dan pertolongan-Nya sehingga penulis dapat menyelesaikan laporan Tugas Akhir dengan judul **“Pemodelan Angka Putus Sekolah Usia SMA di Indonesia dengan Pendekatan Regresi Nonparametrik Spline”**. Dalam penyelesaian tugas akhir ini, penulis mendapatkan banyak bantuan, bimbingan, dan perhatian dari berbagai pihak. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Suhartono selaku Kepala Departemen Statistika ITS
2. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah membimbing, memberikan ilmu dan nasehat dalam menyelesaikan Tugas Akhir ini.
3. Ibu Dr. Vita Ratnasari, S.Si, M.Si dan ibu Dra. Madu Ratna M.Si selaku dosen penguji yang telah memberikan saran, kritik dan masukan demi kesempurnaan Tugas Akhir ini.
4. Kedua orang tua dan keluargayang telah mendoakan memberikan semangat, motivasi serta dukungan selama ini.
5. Bapak Prof. Drs. Nur Iriawan M.IKom. Ph.D. selaku dosen wali saya yang telah membimbing penulis selama mengikuti kuliah di Statistika ITS
6. Serta pihak-pihak lain yang tidak dapat penulis sebutkan satu persatu

Penulis menyadari bahwa Tugas Akhir ini masih belum sempurna, oleh karena itu saran dan kritik yang membangun sangat diharapkan demi kesempurnaan Tugas Akhir ini. Semoga Tugas Akhir ini dapat memberikan manfaat bagi pembaca dan semua pihak.

Surabaya, Juli 2017

Penulis

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA	
2.1 Statistika Deskriptif.....	7
2.2 Analisis Regresi.....	7
2.3 Regresi Nonparametrik Spline	8
2.4 Pemilihan Titik Knot Optimal	9
2.5 Koefisien Determinasi (R^2).....	10
2.6 Pengujian Parameter Model	10
2.6.1 Uji Serentak	11
2.6.2 Uji Parsial	12
2.7 Pemeriksaan Asumsi Residual	12
2.7.1 Asumsi Residual Identik.....	12
2.7.2 Asumsi Residual Independen.....	13
2.7.3 Asumsi Residual Berdistribusi Normal	14
2.8 Angka Putus Sekolah.....	14
2.9 Faktor-Faktor yang Mempengaruhi Angka Putus Sekolah Usia SMA	15
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	19

3.2	Definisi Operasional Variabel	20
3.3	Langkah Analisis dan Diagram Alir	21
BAB IV ANALISIS DAN PEMBAHASAN		
4.1	Karakteristik Angka Putus Sekolah Usia SMA dan Faktor yang Mempengaruhinya.....	23
4.2	Pola Hubungan Angka Putus Sekolah Usia SMA di Indonesia dengan Faktor-Faktor yang diduga Berpengaruh	26
4.3	Pemilihan Titik Knot Optimum.....	29
4.3.1	Pemilihan Titik Knot dengan Satu Titik Knot	30
4.3.2	Pemilihan Titik Knot dengan Dua Titik Knot.....	31
4.3.3	Pemilihan Titik Knot dengan Tiga Titik Knot	32
4.3.4	Pemilihan Titik Knot dengan Kombinasi Titik Knot....	33
4.3.5	Pemilihan Model Terbaik.....	34
4.3.6	Penaksiran Parameter Model Regresi Nonparametrik Spline.....	35
4.4	Pengujian Parameter Model Regresi Nonparametrik Spline	35
4.4.1	Uji Serentak	36
4.4.2	Uji Individu	36
4.5	Pemodelan Rergresi Nonparametrik Spline Optimal dengan Dua Variabel Prediktor	37
4.5.1	Pemilihan Titik Knot dengan Satu Titik Knot	38
4.5.2	Pemilihan Titik Knot dengan Dua Titik Knot.....	38
4.5.3	Pemilihan Titik Knot dengan Tiga Titik Knot	40
4.5.4	Pemilihan Titik Knot dengan Kombinasi Knot.....	41
4.5.5	Pemilihan Model Terbaik.....	42
4.5.6	Penaksiran Parameter Model Regresi Nonparametrik Spilne.....	42
4.6	Pengujian Parameter Model Regresi Nonparametrik Spline	42
4.6.1	Uji Serentak	43
4.6.2	Uji Individu	43
4.7	Pengujian Asumsi Residual	44
4.7.1	Uji Asumsi Identik.....	44

4.7.2 Pemeriksaan Asumsi Independen	45
4.7.3 Uji Asumsi Distribusi Normal	46
4.8 Nilai Koefisien Determinasi (R^2)	47
4.9 Interpretasi Model Regresi Nonparametrik Spline	47
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan.....	51
5.2 Saran.....	52
DAFTAR PUSTAKA	53
LAMPIRAN	55
BIODATA PENULIS	92

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 Analisis Ragam.....	11
Tabel 3.1 Variabel Penelitian	19
Tabel 3.2 Struktur Data Variabel Penelitian.....	20
Tabel 4.1 Karakteristik Angka Putus Sekolah	23
Tabel 4.2 Nilai GCV dengan Satu Titik Knot	30
Tabel 4.3 Nilai GCV dengan Dua Titik Knot	31
Tabel 4.4 Nilai GCV dengan Tiga Titik Knot	32
Tabel 4.5 Nilai GCV dengan Kombinasi Knot.....	34
Tabel 4.6 Perbandingan Nilai GCV dari 1 Titik Knot, 2 Titik Knot, 3 Titik Knot, dan Kombinasi Knot.....	35
Tabel 4.7 Analisis Ragam Uji Serentak.....	36
Tabel 4.8 Hasil Uji Individu	37
Tabel 4.9 Nilai GCV dengan Satu Titik Knot Pada Dua Variabel Prediktor.....	38
Tabel 4.10 Nilai GCV dengan Dua Titik Knot Pada Dua Variabel.....	39
Tabel 4.11 Nilai GCV dengan Tiga Titik Knot Pada Dua Variabel Prediktor.....	40
Tabel 4.12 Nilai GCV dengan Kombinasi Knot Pada Dua Variabel Prediktor.....	41
Tabel 4.13 Nilai GCV Minimum Dari Hasil Titik Knot Optimal Pada Dua Variabel Prediktor	42
Tabel 4.14 Analisis Ragam Uji Serentak Dua Variabel Prediktor	43
Tabel 4.15 Hasil Uji Individu	44
Tabel 4.16 Analisis Ragam Uji <i>Glejser</i>	45

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 3.1	Diagram Alir Penelitian22
Gambar 4.1	Diagram Batang Angka Putus Sekolah Usia SMA25
Gambar 4.2	<i>Scatterplot</i> Persentase Penduduk Miskin dengan Angka Putus Sekolah.....27
Gambar 4.3	<i>Scatterplot</i> Persentase Guru SM/Sederajat terhadap Murid Dengan Angka Putus Sekolah..28
Gambar 4.4	<i>Scatterplot</i> Persentase Jumlah Sekolah SM/Sederajat terhadap Murid Dengan Angka Putus Sekolah28
Gambar 4.5	<i>Scatterplot</i> Rata-Rata Lama Sekolah Penduduk Usia 15 Tahun Ke Atas Dengan Angka Putus Sekolah29
Gambar 4.6	Plot ACF Residual45
Gambar 4.7	Hasil Uji Kolmogorov-Smirnov.....46
Gambar 4.8	Peta Persebaran Angka Putus Sekolah Usia SMA Berdasarkan Persentase Guru SM/Sederajat Terhadap Murid49
Gambar 4.9	Peta Persebaran Angka Putus Sekolah Usia SMA Berdasarkan Rata-Rata Lama Sekolah Penduduk Usia 15 Ke atas50

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Data Angka Putus Sekolah Usia SMA di Indonesia dengan Faktor-Faktor yang Mempengaruhi Tahun 2016.....55
Lampiran 2	Program Pemilihan Titik Knot dengan Satu Titik Knot Menggunakan Software R.....57
Lampiran 3	Program Pemilihan Titik Knot dengan Dua Titik Knot Menggunakan <i>Software R</i>60
Lampiran 4	Program Pemilihan Titik Knot dengan Tiga Titik Knot Menggunakan <i>Software R</i>63
Lampiran 5	Program Pemilihan Titik Knot Empat Variabel dengan Kombinasi Titik Knot Menggunakan <i>Software R</i>66
Lampiran 6	Program Estimasi Parameter dengan Kombinasi Titik Knot72
Lampiran 7	Program Uji Glejser untuk Kombinasi Titik Knot75
Lampiran 8	Output Nilai GCV dengan Satu Titik Knot Pada Empat Variabel Prediktor77
Lampiran 9	Output Nilai GCV dengan Dua Titik Knot Pada Empat Variabel Prediktor78
Lampiran 10	Output Nilai GCV dengan Tiga Titik Knot Pada Empat Variabel Prediktor79
Lampiran 11	Output Nilai GCV dengan Kombinasi Titik Knot Pada Empat Variabel Prediktor80
Lampiran 12	Output Nilai GCV dengan Satu Titik Knot Pada Dua Variabel Prediktor81
Lampiran 13	Output Nilai GCV dengan Dua Titik Knot Pada Dua Variabel Prediktor82
Lampiran 14	Output Nilai GCV dengan Tiga Titik Knot Pada Dua Variabel Prediktor83
Lampiran 15	Output Nilai GCV dengan Kombinasi Titik Knot Pada Dua Variabel Prediktor84

Lampiran 16	Output Estimasi Parameter dan Uji Signifikansi Parameter Model Pada Empat Variabel Prediktor	85
Lampiran 17	Output Estimasi Parameter dan Uji Signifikansi Parameter Model Pada Dua Variabel Prediktor	87
Lampiran 18	Output Uji <i>Glejser</i>	89
Lampiran 19	Surat Pernyataan Data.....	91

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara kepulauan yang terkenal dengan keanekaragamannya. Sebagai negara kepulauan Indonesia memiliki sumber daya alam yang berlimpah. Selain sumber daya alamnya, Indonesia juga kaya akan sumber daya manusianya. Berdasarkan hasil survei penduduk tahun 2010 yang dilakukan oleh Badan Pusat Statistik, jumlah penduduk Indonesia mencapai 237.556.363 jiwa (BPS, 2010). Hal ini menunjukkan Indonesia menempati peringkat ke 4 di dunia dengan jumlah penduduk terbanyak. Jumlah sumber daya manusia di Indonesia yang melimpah harus diimbangi dengan sumber daya manusia yang berkualitas. Salah satu indikator sumber daya manusia yang berkualitas adalah pendidikan.

Pendidikan mempunyai peran yang sangat penting untuk membangun suatu negara. Pemberian pendidikan formal, non formal maupun informal dari usia dini bisa menciptakan sumber daya manusia yang berkualitas pada masa yang akan datang dan diharapkan dapat memberikan kontribusi positif dalam berbagai aspek kehidupan untuk kemajuan negara. Indonesia merupakan salah satu negara yang berusaha untuk meningkatkan taraf pendidikan penduduknya untuk menciptakan sumber daya manusia yang berkualitas. Pemerintah telah melakukan berbagai program untuk meningkatkan taraf pendidikan penduduk Indonesia, salah satunya dengan melakukan Pelaksanaan Program Indonesia Pintar (PIP) melalui usia wajib belajar 12 tahun. Melalui program ini pemerintah berupaya mencegah peserta didik dari kemungkinan putus sekolah, dan diharapkan dapat menarik siswa putus sekolah agar kembali melanjutkan pendidikannya (Kemendikbud, 2015).Berlawanan dengan harapan yang ada, kenyataannya angka putus sekolah di Indonesia pada seluruh jenjang pendidikan masih tinggi, terutama pada jenjang pendidikan usia SMA.

Tingginya angka putus sekolah usia SMA ditunjukkan dengan laporan *Education at Glance 2015* yang dikeluarkan oleh *Organisation for Economic Cooperation and Development* (OECD) yaitu Indonesia pernah menempati urutan ke 2 di dunia dengan angka putus sekolah jenjang SMA tertinggi yaitu mencapai 60% dari masyarakat kelompok usia 25-34 tahun yang gagal menyelesaikan pendidikan menengah. Pada tahun 2016 angka putus sekolah atau persentase yang tidak bersekolah pada usia SMA di Indonesia mencapai 4,98 persen. Angka ini cukup tinggi bila dibandingkan dengan kelompok umur lainnya. Angka putus sekolah kelompok umur 7-12 tahun yaitu sebesar 0,76 persen, sedangkan untuk usia 13-15 tahun adalah sebesar 4,60 persen. Hal ini juga dapat dilihat pada jumlah anak usia SMA yang bersekolah sebesar 8.433.971 anak. Sementara itu, angka partisipasi untuk usia SMA sebesar 70,83 persen. Hal ini dapat dikatakan bahwa hanya 71 persen saja anak usia SMA yang bisa sekolah, sisanya 29 persen tidak bisa bersekolah di jenjang tersebut atau terdapat 2.445.852 anak usia SMA tidak bisa bersekolah di jenjang SMA (BPS, 2016). Akibat dari banyaknya anak yang tidak bisa bersekolah antara lain semakin meningkatnya jumlah penduduk yang berada di bawah garis kemiskinan dan banyaknya jumlah pengangguran, sehingga akan menimbulkan permasalahan sosial, kesehatan, dan meningkatnya tingkat kriminalitas. Tingginya angka putus sekolah usia SMA di Indonesia maka perlu dilakukan penelitian mengenai faktor-faktor yang mempengaruhi angka putus sekolah usia SMA di Indonesia.

Upaya untuk menekan angka putus sekolah usia SMA di Indonesia, perlu dipahami faktor apa saja yang mempengaruhinya agar dapat dianalisis. Salah satu metode yang dapat digunakan menganalisis faktor-faktor yang mempengaruhi angka putus sekolah usia SMA adalah analisis regresi. Analisis regresi merupakan suatu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor (Draper & Smith, 1992). Metode analisis regresi dibagi menjadi 3 jenis yaitu regresi parametrik, regresi nonparametrik,

dan regresi semiparametrik. Regresi parametrik digunakan jika pola data membentuk pola tertentu seperti kubik, linier atau kuadratik. Regresi nonparametrik digunakan jika pola data tidak membentuk pola tertentu, sedangkan regresi semiparametrik digunakan jika sebagian pola data membentuk pola dan sebagian lagi pola data tidak membentuk pola tertentu (Budiantara, 2007).

Pada penelitian ini menggunakan regresi nonparametrik dikarenakan pola hubungan antara angka putus sekolah usia SMA dengan faktor-faktor yang diduga mempengaruhinya tidak membentuk pola tertentu. Fungsi yang dapat digunakan dalam regresi nonparametrik pada penelitian ini adalah spline, dikarenakan spline merupakan potongan polinomial yang mempunyai sifat tersegmen, sehingga mempunyai fleksibilitas tinggi dan dapat menyesuaikan diri terhadap karakteristik lokal suatu data. Penggunaan analisis regresi nonparametrik spline sangat baik digunakan untuk mengatasi pola data yang memiliki perubahan perilaku pada sub-sub interval tertentu dan mempunyai kemampuan yang sangat baik untuk digeneralisasikan pada pemodelan statistika yang kompleks dan rumit (Budiantara, 2009).

Beberapa penelitian angka putus sekolah yang pernah dilakukan yaitu Citrasari (2011) yang berjudul pemodelan angka putus sekolah bagi anak usia wajib belajar di Jawa Timur dengan pendekatan *Generalized Poisson Regression*. Penelitian berikutnya pernah dilakukan oleh Pradipta (2013) mengenai pemodelan angka putus sekolah usia SMA di Jawa timur dengan pendekatan spline multivariabel. Penelitian selanjutnya oleh Mubarokah (2016) melakukan penelitian mengenai pemodelan angka putus sekolah usia SMP menggunakan metode regresi nonparametrik spline di Papua. Berdasarkan penelitian yang dilakukan Citrasari (2011), Pradipta (2013) dan Mubarokah (2016) menyimpulkan bahwa faktor-faktor yang mempengaruhi angka putus sekolah yaitu laju pertumbuhan ekonomi, persentase penduduk miskin, persentase guru terhadap murid, dan persentase sekolah terhadap murid. Penelitian selanjutnya oleh Septiana (2011) mengenai pemodelan remaja putus sekolah usia SMA di

Provinsi Jawa Timur menggunakan metode regresi spasial yang menyimpulkan bahwa faktor yang berpengaruh signifikan terhadap angka putus sekolah adalah keluarga miskin dan didapatkan hasil nilai R^2 yang diperoleh sangat rendah yaitu kurang dari 50 persen. Oleh karena itu, perlu dilakukan metode yang dapat menghasilkan nilai R^2 yang lebih baik.

Berdasarkan uraian latar belakang diatas, maka dilakukan penelitian mengenai “Pemodelan Angka Putus Sekolah Usia SMA di Indonesia dengan Pendekatan Regresi Nonparametrik Spline”. Dari hasil penelitian ini diharapkan dapat memberikan informasi tambahan terhadap pemerintah untuk merencanakan program yang bertujuan menurunkan angka putus sekolah khususnya usia SMA di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik angka putus sekolah usia SMA dan faktor-faktor yang diduga mempengaruhinya?
2. Bagaimana pemodelan angka putus sekolah usia SMA di Indonesia menggunakan pendekatan regresi nonparametrik spline?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik angka putus sekolah usia SMA dan faktor-faktor yang diduga mempengaruhinya.
2. Memodelkan angka putus sekolah usia SMA di Indonesia dengan menggunakan pendekatan regresi nonparametrik spline.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Manfaat untuk Pemerintah
Hasil dari analisis dan pemodelan dalam penelitian ini dapat digunakan sebagai bahan evaluasi yang dapat digunakan pemerintah dalam menurunkan angka putus sekolah usia SMA di Indonesia.
2. Manfaat untuk Bidang Keilmuan
Hasil penelitian ini diharapkan dapat memberikan informasi dan pengetahuan mengenai penerapan metode regresi nonparamterik spline pada pemodelan angka putus sekolah usia SMA di Indonesia sehingga dapat dijadikan referensi untuk penelitian selanjutnya.

1.5 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Data yang digunakan dalam penelitian ini diambil dari Badan Pusat Statistik tahun 2016.
2. Fungsi spline yang digunakan adalah *SplineTruncated Linear*.
3. Titik knot yang digunakan adalah satu, dua, tiga, dan kombinasi knot.
4. Pemilihan titik knot optimal menggunakan metode *Generalized Cross Validation (GCV)*.

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna (Wapole, 1995). Statistika deskriptif meliputi pengukuran pemusatan data, pengukuran penyebaran data serta membuat dan menampilkan grafik atau diagram. Ukuran pemusatan data biasanya dihitung dari nilai rata-rata (*mean*), sedangkan ukuran penyebaran data diukur oleh nilai varians. Selain itu, nilai minimum dan maksimum pada suatu data juga sering digunakan dalam analisis statistika deskriptif. Berikut ini adalah rumus yang digunakan untuk menghitung *mean* :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

dengan

\bar{x} : *Mean*

x_i : pengamatan ke- i , $i = 1, 2, \dots, n$

n : banyaknya pengamatan

Varian (s^2) adalah kuadrat simpangan dari semua nilai data terhadap rata-rata hitung, dengan rumus sebagai berikut.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (2.2)$$

2.2 Analisis Regresi

Menurut Draper dan Smith (1992) analisis regresi merupakan sebuah metode statistika yang memberikan penjelasan tentang pola hubungan antara dua variabel atau lebih. Variabel tersebut terdiri dua jenis yaitu variabel respon dan variabel

prediktor. Jika dimisalkan variabel respon adalah y dan variabelprediktor adalah x maka model regresi diberikan oleh.

$$y_i = f(x_i) + \varepsilon_i \quad (2.3)$$

dengan f adalah kurva regresi dan ε_i adalah error yang diasumsikan identik, independen, berdistribusi normal dengan *mean* nol dan variansi σ^2 (Budiantara, 2009). Dalam analisis regresi sering kali didahului dengan mempelajari pola hubungan antar variabel respon dan variabel prediktor menggunakan *scatter plot*. Berdasarkan *plot* tersebut, bentuk kurva dapat diketahui polanya apakah linier, kuadratik, atau kubik. Tetapi dapat juga kurva yang dihasilkan tidak memiliki bentuk pola tertentu. Oleh karena itu dalam analisis regresi terdapat tiga pendekatan yaitu regresi parametrik, nonparametrik, dan semiparametrik. Analisis regresi parametrik digunakan apabila pola hubungan antara variabel respon dan variabel prediktor yang dihasilkan membentuk suatu pola. Analisis regresi nonparametrik digunakan apabila pola hubungan antara variabel respon dan variabel prediktor tidak membentuk suatu pola. Analisis regresi semiparametrik digunakan apabila pola hubungan antara variabel respon dan variabel prediktor sebagian ada yang membentuk pola dan sebagian lagi ada yang tidak membentuk pola.

2.3 Regresi Nonparametrik Spline

Regresi nonparametrik merupakan metode yang dapat digunakan ketika pola antara variabel respon dan variabel prediktor tidak diketahui bentuknya. Pemodelan dalam statistika diharapkan memiliki model yang sederhana seperti regresi parametrik. Namun pada kenyataannya terdapat model untuk sekumpulan data memiliki pola yang kompleks. Sehingga, jika model sederhana dipaksakan maka akan menghasilkan *error* yang besar (Budiantara, 2009). Untuk mengatasi hal ini, perlu digunakan pendekatan dengan regresi nonparametrik. Regresi nonparametrik bersifat fleksibel artinya data dapat mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh

subjektifitas (Eubank, 1988). Spline memiliki fleksibilitas yang tinggi, mampu menangani data atau fungsi yang mulus (*smooth*), dan memiliki kemampuan yang sangat baik dalam menangani data yang perilakunya berubah-ubah dalam sub interval tertentu. Model regresi nonparametrik secara umum dapat disajikan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i \quad ; i = 1, 2, \dots, n \quad (2.4)$$

Fungsi $f(x_i)$ merupakan fungsi regresi yang bentuk polanya tidak diketahui dan $\varepsilon_i \sim \text{IIDN}(0, \sigma^2)$. Jika $f(x_i)$ adalah fungsi spline berorde m yang merupakan derajat polinomial dan dengan titik knot k_1, k_2, \dots, k_r maka $f(x_i)$ dapat disajikan menjadi

$$f(x_i) = \sum_{j=0}^m \beta_j x_i^j + \sum_{q=1}^r \beta_{m+q} (x_i - k_q)_+^m \quad (2.5)$$

Apabila persamaan (2.5) disubstitusikan ke dalam persamaan (2.4) maka akan diperoleh persamaan regresi nonparametrik splines sebagai berikut.

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{q=1}^r \beta_{m+q} (x_i - k_q)_+^m + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.6)$$

Fungsi $(x_i - k_q)_+^m$ merupakan fungsi potongan yang diberikan oleh

$$(x_i - k_q)_+^m = \begin{cases} (x_i - k_q)^m, & x_i \geq k_q \\ 0, & x_i < k_q \end{cases} \quad (2.7)$$

2.4 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan perilaku fungsi pada interval yang berlainan. Oleh karena itu, untuk mendapatkan model regresi spline terbaik yang paling sesuai dengan data perlu dicari nilai knot optimal. Salah satu metode yang banyak dipakai untuk menentukan titik knot optimal adalah *Generalized Cross Validation* (GCV). Titik knot yang optimal didapatkan dari nilai GCV yang paling

minimum. Fungsi GCV didefinisikan sebagai berikut (Eubank, 1988).

$$GCV(k_1, k_2, \dots, k_r) = \frac{MSE(k_1, k_2, \dots, k_r)}{\left(n^{-1} \text{trace}[\mathbf{I} - \mathbf{A}(k_1, k_2, \dots, k_r)]\right)^2} \quad (2.8)$$

dimana n adalah banyak pengamatan, \mathbf{I} adalah matrik identitas,

$$\mathbf{A}(k_1, k_2, \dots, k_r) = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$$

$MSE(k_1, k_2, \dots, k_r)$ adalah sebagai berikut

$$MSE(k_1, k_2, \dots, k_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.9)$$

2.5 Koefisien Determinasi (R^2)

Semakin tinggi nilai R^2 yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model tersebut dalam menjelaskan variabel respon (Draper & Smith, 1992). Berikut ini merupakan rumus untuk menghitung R^2 .

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.10)$$

Selain itu, pemilihan model juga akan memperhatikan banyak parameter yang digunakan dalam model tersebut. Hal ini dijelaskan oleh prinsip parsimoni, dimana suatu model regresi yang baik adalah model regresi dengan jumlah parameter sesedikit mungkin tetapi menghasilkan R^2 yang cukup tinggi.

2.6 Pengujian Parameter Model

Pengujian parameter dilakukan untuk mengetahui apakah suatu variabel memberikan pengaruh secara signifikan terhadap model. Pada regresi spline, uji parameter model dilakukan setelah mendapatkan model regresi dengan titik knot optimal berdasarkan nilai GCV yang paling minimum. Pengujian parameter dilakukan

dengan dua tahapan yaitu pengujian secara serentak dan individu (parsial) dengan penjelasan sebagai berikut.

2.6.1 Uji Serentak

Uji serentak adalah uji signifikansi model secara keseluruhan atau untuk mengetahui apakah semua variabel prediktor yang dimasukkan ke dalam model memberikan pengaruh secara bersama-sama. Model yang digunakan adalah sebagai berikut.

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{q=1}^r \beta_{m+q} \left(x_i - k_q \right)_+^m + \varepsilon_i, i = 1, 2, \dots, n$$

Berdasarkan persamaan model di atas, maka hipotesis yang digunakan untuk pengujian secara serentak adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_m = \beta_{m+1} = \dots = \beta_{m+r} = 0$$

$$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, m+r$$

Nilai $m+r$ adalah banyak parameter dalam model regresi dan n adalah jumlah observasi. Berikut ini merupakan statistik uji yang digunakan.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (2.11)$$

Perhitungan nilai statistik uji F didapatkan dari Analisis Ragam (ANOVA) sebagaimana yang ditunjukkan Tabel 2.1

Tabel 2. 1 Analisis Ragam

Sumber Variasi	df	Sum of Square	Mean Square	F hitung
Regresi	$m + r$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$\frac{MS_{regresi}}{MS_{residual}}$
Error	$n - (m + r) - 1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{SS_{residual}}{df_{residual}}$	—
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	—	—

2.6.2 Uji Parsial

Uji parsial digunakan untuk mengetahui variabel yang memberikan pengaruh signifikan terhadap model. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, m + r.$$

Statistik uji:

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} ; j = 1, 2, \dots, m + r \quad (2.12)$$

Dengan

$\hat{\beta}_j$: penaksir parameter β_j

$$SE(\hat{\beta}_j) = \sqrt{\text{diag}[(\mathbf{X}'\mathbf{X})^{-1}\hat{\sigma}^2]_j}$$

Daerah penolakan yang digunakan adalah dengan membandingkan nilai t dengan $t_{\frac{\alpha}{2}, n-(m+r)-1}$. Keputusan tolak H_0 jika $|t| > t_{\frac{\alpha}{2}, n-(m+r)-1}$ yang berarti bahwa parameter signifikan terhadap model.

2.7 Pemeriksaan Asumsi Residual

Uji asumsi residual dilakukan untuk mengetahui kelayakan suatu model regresi yang dihasilkan. Uji asumsi dilakukan terhadap residual yang dihasilkan dari model regresi. Uji asumsi yang harus dipenuhi adalah uji asumsi identik, independen, dan berdistribusi normal.

2.7.1 Asumsi Residual Identik

Uji asumsi identik digunakan untuk mengetahui homogenitas variasi residual. Jika asumsi tidak terpenuhi artinya terdapat heteroskedastisitas yang mengakibatkan kerugian bagi estimator (Eubank, 1988). Salah satu cara untuk mendeteksi adanya heteroskedastisitas adalah menggunakan uji *Glejser*. Uji *Glejser* dilakukan dengan meregresikan *absolut* dari residual dengan variabel prediktornya (Gujarati, 2003). Hipotesis yang digunakan untuk uji *Glejser* adalah sebagai berikut.

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1: \text{minimal ada satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$$

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{\left[\sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 \right] / (n - (m + r))}{\left[\sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2 \right] / ((m + r) - 1)} \quad (2.13)$$

Pengambilan keputusan dari uji *Glejser* adalah tolak H_0 , jika

$F_{hitung} > F_{tabel} \left(F_{\alpha; ((m+r)-1, n-(m+r))} \right)$ atau tolak H_0 apabila $P_{value} < \alpha$ yang berarti terdapat indikasi adanya kasus heteroskedastisitas.

2.7.2 Asumsi Residual Independen

Asumsi selanjutnya yang harus dipenuhi adalah tidak terdapat korelasi antar residual atau autokorelasi. Pendeteksian autokorelasi dapat dilakukan dengan membuat plot *Autocorrelation Function* (ACF). Berikut ini merupakan rumus untuk menghitung ACF (Wei, 2006).

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (2.14)$$

Apabila nilai $\hat{\rho}_k$ keluar dari batas signifikan maka dapat dikatakan terdapat kasus autokorelasi. Sebaliknya jika tidak terdapat nilai $\hat{\rho}_k$ yang keluar dari batas signifikan maka tidak terdapat kasus autokorelasi. Interval konfidensi $(1 - \alpha)100\%$ untuk autokorelasi ρ_k diberikan sebagai berikut.

$$-t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_k) \leq \rho_k \leq t_{n-1; \frac{\alpha}{2}} SE(\hat{\rho}_k). \quad (2.15)$$

$$\text{dimana } SE(\hat{\rho}_k) = \sqrt{\frac{1 + 2 \sum_{l=1}^{k-1} \hat{\rho}_k^2}{n}}$$

2.7.3 Asumsi Residual Berdistribusi Normal

Residual dari suatu model regresi harus mengikuti distribusi normal dengan nilai mean nol dan varians σ^2 . Pengujian distribusi normal secara visual bisa dilakukan dengan *normal probability plot residual*. Residual berdistribusi normal apabila *plot* cenderung mengikuti garis lurus 45° . Cara lain yang sering digunakan adalah pengujian distribusi normal *Kolmogorov-Smirnov* (Daniel, 1989). Berikut ini merupakan hipotesis untuk uji *Kolmogorov-Smirnov*.

$H_0 : F(x) = F_0(x)$ (Residual berdistribusi Normal)

$H_1 : F(x) \neq F_0(x)$ (Residual tidak berdistribusi Normal)

$$D = \max |S_N(x) - F_0(x)| \quad (2.16)$$

dimana

$F_0(x)$ = fungsi distribusi frekuensi kumulatif

$S_N(x)$ = fungsi peluang kumulatif yang diobservasi dari suatu sampel random

Daerah penolakan H_0 adalah apabila $|D| > q_{(1-\alpha)}$ dengan nilai $q_{(1-\alpha)}$ didapatkan dari tabel *Kolmogorov-Smirnov*.

2.8 Angka Putus Sekolah

Putus sekolah merupakan suatu kondisi dimana anak berhenti secara terpaksa dari suatu lembaga pendidikan tempat dia belajar. Sehingga dapat diartikan bahwa putus sekolah merupakan suatu kondisi dimana terlantarnya anak dari suatu lembaga pendidikan formal yang disebabkan beberapa faktor (Musfion, 2009). Angka putus sekolah didapatkan dari proporsi anak menurut usia kelompok usia sekolah yang sudah tidak bersekolah lagi atau tidak menamatkan suatu jenjang tertentu.

Kelompok umur pada penelitian ini yang dimaksud adalah kelompok usia SMA yaitu 16-18 tahun (Sirusa BPS, 2017).

$$APTS \text{ (16-18 tahun)} = \frac{P_{16-18(PS)}}{P_{16-18}} \times 100\% \quad (2.17)$$

Dimana

APTS : Angka Putus Sekolah

PS : Putus Sekolah

$P_{16-18(PS)}$: Jumlah penduduk usia 16-18 tahun yang tidak bersekolah lagi

P_{16-18} : Jumlah penduduk usia 16-18 tahun yang pernah/sedang bersekolah

2.9 Faktor-Faktor yang Mempengaruhi Angka Putus Sekolah Usia SMA

Menurut Badan Pusat Statistik (2015) angka putus sekolah disebabkan beberapa yaitu faktor ekonomi, sosial dan budaya, dan geografis. Berikut ini akan jelaskan aspek-aspek mengenai faktor-faktor penyebab angka putus sekolah.

1. Faktor Ekonomi

Berdasarkan BPS, faktor ketidakmampuan membiayai sekolah atau faktor ekonomi menjadi faktor penyebab yang paling dominan dari anak tidak bersekolah. Menurut Beebey (1982) putus sekolah lebih merupakan masalah sosial-ekonomi dari pada masalah pendidikan, karena mayoritas penyebab umum terjadinya putus sekolah adalah kemiskinan atau kesulitan ekonomi. Selain itu, menurut Pradipta (2013) mengenai pemodelan angka putus sekolah usia SMA di Jawa timur dengan pendekatan splinemultivariabel memberikan kesimpulan bahwa faktor faktor-faktor yang mempengaruhi angka putus sekolah yaitu persentase penduduk miskin, persentase guru terhadap murid, dan persentase sekolah terhadap murid. Pada umumnya timbulnya masalah anak putus sekolah berkaitan dengan masalah kemiskinan. Keadaan ekonomi keluarga yang kurang mendukung maka ada banyak

anak usia sekolah harus kehilangan kesempatan memperoleh pendidikan.

2. Faktor Sosial dan Budaya

Menurut Musfiquon (2009) sosiologis anak juga memiliki pengaruh terhadap putus sekolah. Anak yang berada dalam lingkungan keluarga atau masyarakat yang berpendidikan rendah cenderung tidak bersekolah karena pengaruh lingkungan disekitarnya. Berdasarkan penelitian yang dilakukan oleh Farah (2014) mengenai faktor penyebab putus sekolah dan dampak negatifnya bagi anak, memberikan kesimpulan bahwa salah satu faktor penyebab putus sekolah adalah faktor sosial/budaya dikarenakan pengaruh pola pikir dan budaya orang tua/masyarakat yang tidak begitu mementingkan pendidikan anak.

3. Faktor Geografis

Distribusi lokasi sangat mempengaruhi kemungkinan anak putus sekolah. Menurut Deputy Evaluasi Kerja Pembangunan Badan Perencanaan Pembangunan Nasional (2009), Angka Partisipasi Kasar (APK) dan Angka Partisipasi Murni (APM) dipengaruhi oleh persentase guru terhadap murid dan persentase jumlah sekolah terhadap murid. Hipotesis yang digunakan yaitu semakin banyaknya guru yang tersedia akan meningkatkan APK dan APM. Sedangkan hubungan antara persentase jumlah sekolah terhadap dengan APK dan APM diharapkan positif yang diketahui bahwa semakin banyak sekolah yang tersedia akan meningkatkan APK dan APM. Hal ini ditunjukkan dengan penelitian yang telah dilakukan oleh Citrasari (2011) yang meneliti mengenai angka putus sekolah bagi anak usia wajib belajar di Jawa Timur dengan pendekatan *Generalized Poisson Regresion* yang dihasilkan bahwa persentase laju pertumbuhan ekonomi, persentase guru (SD/MI dan SMP/MTS) terhadap siswa dan tingkat kesempatan kerja di Jawa Timur berpengaruh signifikan terhadap angka putus sekolah. Selain penelitian yang dilakukan citrasari, penelitian selanjutnya dilakukan oleh Pradipta (2013) mengenai pemodelan angka putus sekolah usia SMA di Jawa timur dengan pendekatan spline multivariabel memberikan

kesimpulan bahwa faktor faktor-faktor yang mempengaruhi angka putus sekolah yaitu persentase penduduk miskin, persentase guru terhadap murid, dan persentase jumlah sekolah terhadap murid.

Berdasarkan studi kepustakaan diatas dapat ditarik kesimpulan suatu informasi bahwa variabel yang berpengaruh terhadap angka putus sekolah adalah persentase penduduk miskin (X_1), persentase guru terhadap murid (X_2), persentase jumlah sekolah terhadap murid (X_3), dan rata-rata lama sekolah usia 15 tahun ke atas (X_4). Sehingga, variabel tersebut digunakan dalam penelitian tentang angka putus sekolah usia SMA di Indonesia dengan pendekatan regresi nonparametrik spline.

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder, yang didapat dari Badan Pusat Statistik (BPS) tahun 2016. Data tersebut mencakup tentang angka putus sekolah jenjang SMA dan faktor-faktor yang diduga mempengaruhinya di Indonesia. Unit penelitian yang digunakan adalah 34 provinsi yang ada di Indonesia.

3.1 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini ada 2 yaitu variabel respon dan variabel prediktor. Variabel respon yang digunakan dalam penelitian ini adalah angka putus sekolah usia SMA di Indonesia dengan 4 variabel yang prediktor yang diduga mempengaruhi angka putus sekolah usia SMA di Indonesia. Variabel-variabel yang digunakan dalam penelitian ini tercantum Tabel 3.1.

Tabel 3. 1 Variabel Penelitian

Variabel	Keterangan	Skala
Y	Angka Putus Sekolah Usia SMA di Indonesia	Rasio
X ₁	Persentase penduduk miskin	Rasio
X ₂	Persentase guru SM/ sederajat terhadap murid	Rasio
X ₃	Persentase jumlah sekolah SM/ sederajat terhadap murid	Rasio
X ₄	Rata-rata lama sekolah penduduk usia 15 ke atas	Rasio

Selanjutnya struktur data yang digunakan dalam penelitian disajikan pada Tabel 3.2.

Tabel 3. 2Struktur Data Variabel Penelitian

Provinsi	y	x ₁	x ₂	x ₃	x ₄
1	y ₁	x _{1,1}	x _{2,1}	x _{3,1}	x _{4,1}
2	y ₂	x _{1,2}	x _{2,2}	x _{3,2}	x _{4,2}
3	y ₃	x _{1,3}	x _{2,3}	x _{3,3}	x _{4,3}
.
.
.
34	y ₃₄	x _{1,34}	x _{2,34}	x _{3,34}	x _{4,34}

3.2 Definisi Operasional Variabel

Definisi operasional variabel-variabel yang digunakan dalam penelitian ini adalah sebagai berikut.

- Variabel Y menyatakan angka putus sekolah usia SMA. TB Angka putus sekolah usia SMA adalah Proporsi anak menurut kelompok usia sekolah yang sudah tidak bersekolah lagi atau yang tidak menamatkan suatu jenjang pendidikan tertentu. Adapun kelompok umur yang dimaksud 16-18 tahun.
- Variabel X_1 menyatakan Persentase penduduk miskin. Penduduk miskin merupakan persentase penduduk yang memiliki ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Jika penduduk memiliki rata-rata pengeluaran perkapita pengeluaran di bawah garis kemiskinan tergolong dalam penduduk miskin.
- Variabel X_2 menyatakan persentase guru SM/Sederajat terhadap murid. Jumlah guru pendidikan menengah akhir di Indonesia pada masing-masing provinsi dibagi dengan jumlah murid pendidikan menengah akhir di provinsi tersebut dikali 100%.
- Variabel X_3 menyatakan persentase jumlah sekolah SM/Sederajat terhadap murid. Jumlah sekolah tingkat pendidikan sekolah menengah akhir di Indonesia pada masing-masing provinsi dibagi dengan jumlah murid

pendidikan sekolah menengah akhir di provinsi tersebut dikali 100%.

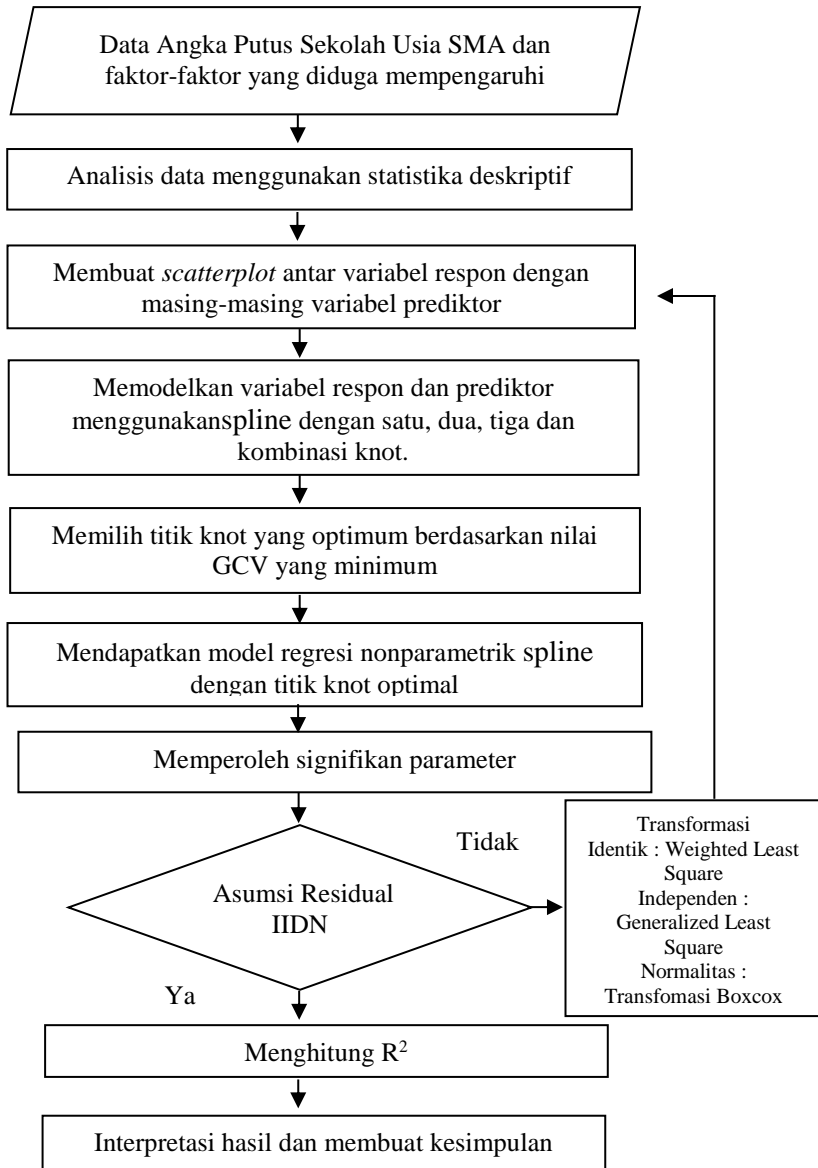
- e. Variabel X_4 menyatakan rata-rata lama sekolah penduduk usia 15 tahun ke atas. Rata-rata lama sekolah usia 15 tahun ke atas adalah jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal minimal 12 tahun (tidak termasuk tahun yang mengulang).

3.3 Langkah Analisis dan Diagram Alir

Langkah analisis yang digunakan dalam menyelesaikan permasalahan dalam penelitian ini adalah sebagai berikut.

1. Melakukan analisis deskriptif tentang angka putus sekolah usia SMA di Indonesia dan karakteristik faktor-faktor yang diduga berpengaruh terhadap angka putus sekolah usia SMA di Indonesia.
2. Membuat *scatterplot* antara variabel respon angka putus sekolah usia SMA di Indonesia dengan masing-masing variabel prediktor yang diduga berpengaruh untuk mengetahui bentuk pola data.
3. Memodelkan variabel respon dan prediktor menggunakan Regresi Nonparametrik Splinedengan satu, dua, tiga dan kombinasi knot.
4. Memilih titik knot optimal berdasarkan nilai GCV yang paling minimum.
5. Mendapatkan model regresi nonparametrik spline dengan titik knot optimal.
6. Menguji signifikansi parameter model regresi nonparametrik spline secara serentak dan parsial.
7. Menguji asumsi residual IIDN dari model regresi nonparametrik splineterbaik.
8. Menghitung nilai koefisien determinasi (R^2).
9. Menginterpretasi model dan menarik kesimpulan.

Berdasarkan langkah-langkah analisis di atas, maka dapat disajikan secara ringkas dalam diagram alir sebagai berikut.



Gambar 3. 1 Diagram Alir Penelitian

BAB IV

ANALISIS DAN PEMBAHASAN

Bab ini membahas tentang karakteristik angka putus sekolah usia SMA di Indonesia tahun 2016 serta faktor-faktor yang diduga berpengaruh menggunakan statistika deskriptif. Selain itu, bab ini juga membahas mengenai pemodelan angka putus sekolah usia SMA menggunakan regresi nonparametrik spline, dimana kurva regresi nonparametrik diperoleh menggunakan fungsi spline linear satu, dua, tiga knot, dan kombinasi knot sehingga dapat diketahui faktor-faktor yang berpengaruh signifikan terhadap angka putus sekolah usia SMA di Indonesia

4.1 Karakteristik Angka Putus Sekolah Usia SMA dan Faktor yang Mempengaruhinya

Semakin tinggi angka putus sekolah maka menunjukkan semakin rendah kualitas suatu program pendidikan dalam meningkatkan sumber daya manusia yang berkualitas. Sehingga hal tersebut akan berpengaruh terhadap kemajuan suatu negara. Karakteristik Angka Putus Sekolah usia SMA dan faktor-faktor yang diduga mempengaruhinya ditunjukkan pada Tabel 4.1 berikut.

Tabel 4. 1Karakteristik Angka Putus Sekolah

Variabel	Mean	Varian	Minimum	Maksimum
Y	5,98	4,67	2,8	10,97
X ₁	11,36	37,57	3,75	28,4
X ₂	7,02	1,57	4,66	10,19
X ₃	0,33	0,01	0,19	0,60
X ₄	8,64	0,77	6,48	10,92

Tabel 4.1 menunjukkan bahwa nilai rata-rata angka putus sekolah usia SMA di Indonesia (Y) adalah sebesar 5,98 persen, dengan varian sebesar 4,67 persen, nilai angka putus sekolah usia SMA paling kecil adalah 2,8 persen yang terdapat pada provinsi DI Yogyakarta. Sementara nilai paling tinggi terdapat pada

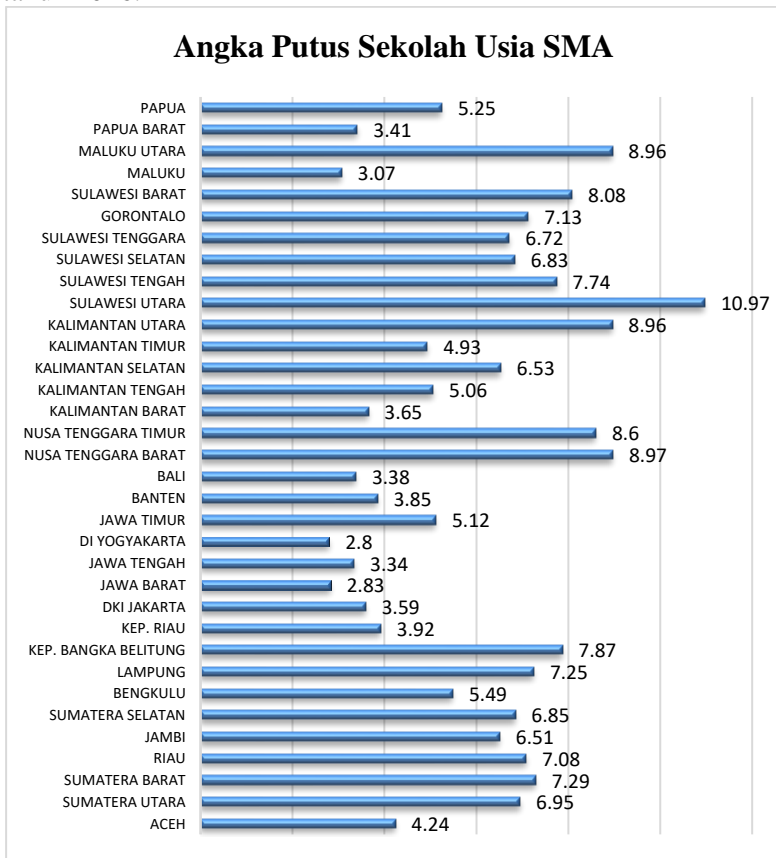
provinsi Sulawesi Sulawesi Utara yaitu mencapai 10,97 persen. Tingginya angka putus sekolah usia SMA di Sulawesi Utara dikarenakan Sulawesi Utara mempunyai daerah pedesaan yang lebih banyak dari pada perkotaan, dimana daerah pedesaan untuk angka putus sekolah lebih tinggi dibandingkan dengan daerah perkotaan. Dari analisis diatas maka provinsi dengan nilai persentase anak putus sekolah tertinggi yaitu provinsi Sulawesi Utara harus lebih meningkatkan mutu program pendidikan agar persentase angka putus sekolah usia SMA dapat berkurang.

Karakteristik variabel X_1 yang diduga mempengaruhi angka putus sekolah usia SMA yaitu persentase penduduk miskin. Dari analisis statistika deskriptif Tabel 4.1 dapat diketahui bahwa variabel persentase penduduk miskin memiliki rata-rata sebesar 11,36 persen dengan varians sebesar 37,57 persen. Nilai varians yang cukup besar mengindikasikan perbedaan nilai maksimum dan nilai minimum variabel X_1 yang besar. Provinsi dengan nilai persentase penduduk miskin terbesar berada pada provinsi Papua yaitu sebesar 28,40 persen dan yang terendah berada pada provinsi DKI Jakarta dengan persentase sebesar 3,75 persen. Variabel X_2 yaitu persentase guru SM/Sederajat terhadap murid. persentase guru SM/Sederajat terhadap murid memiliki rata-rata sebesar 7,02 persen dengan varian sebesar 1,57 persen. Nilai persentase guru SM/Sederajat terhadap murid terkecil berada pada provinsi Banten dengan persentase sebesar 4,66 persen. Persentase terbesar dari persentase guru SM/Sederajat terhadap murid berada pada provinsi Aceh dengan nilai persentase sebesar 10,19 persen.

Variabel X_3 yaitu persentase jumlah sekolah SM/Sederajat terhadap murid. Rata-rata dari persentase jumlah sekolah SM/Sederajat terhadap murid sebesar 0,33 persen dan varian sebesar 0,01. Nilai persentase jumlah sekolah SM/Sederajat terhadap murid terkecil berada di provinsi Bali yaitu 0,19 persen. Sedangkan persentase jumlah sekolah SM/Sederajat terbesar berada pada provinsi Maluku Utara yaitu 0,60 persen. Variabel X_4 yaitu rata-rata lama sekolah penduduk usia 15 keatas.

Berdasarkan Tabel 4.1 variabel tersebut memiliki rata-rata sebesar 8,64persen dengan varians sebesar 0,77 persen. Provinsi yang memiliki rata-rata lama sekolah penduduk usia 15 tahun ke atas terkecil adalah provinsi Papua. Sedangkan provinsi yang memiliki rata-rata lama sekolah penduduk sekolah usia 15 tahun keatas terbesar adalah provinsi DKI Jakarta.

Berikut merupakan diagram batang yang menunjukkan angka putus sekolah usia SMA di setiap provinsi di Indonesia pada tahun 2016.



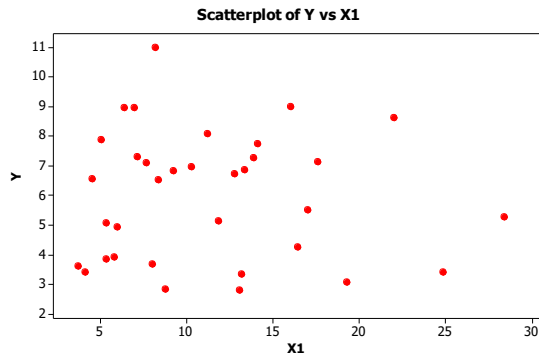
Gambar 4. 1Diagram Batang Angka Putus Sekolah Usia SMA

Berdasarkan Gambar 4.1 menunjukkan provinsi yang memiliki angka putus sekolah usia SMA tertinggi di Indonesia adalah provinsi Sulawesi Utara yaitu mencapai 10,97 persen. Untuk provinsi yang memiliki angka putus sekolah usia SMA terkecil adalah provinsi DI Yogyakarta dengan angka putus sekolah usia SMA mencapai 2,8 persen. Pada gambar 4.1 dapat diketahui bahwa provinsi DKI Jakarta yang merupakan ibukota Indonesia memiliki angka putus sekolah lebih rendah dari rata-rata yaitu mencapai 3,59 persen. Namun, angka putus sekolah usia SMA di DKI Jakarta masih lebih tinggi dari pada provinsi Maluku, Bali, DI Yogyakarta, Jawa Tengah, dan Jawa Barat. Provinsi yang memiliki persentase angka putus sekolah lebih tinggi dari rata-rata adalah provinsi Maluku Utara, Sulawesi Barat, Gorontalo, Sulawesi Tenggara, Sulawesi Selatan, Sulawesi Tengah, Sulawesi Utara, Kalimantan Utara, Kalimantan Selatan, Nusa Tenggara Timur, Nusa Tenggara Barat, Kep Bangka Belitung, Lampung, Sumatera Selatan, Jambi, Riau, Sumatera Barat, Sumatera Utara dan Aceh. Dalam hal ini terlihat provinsi yang ada di Indonesia masih memiliki angka putus sekolah usia SMA yang lebih dari rata-rata, sehingga perlu meningkatkan program pendidikan untuk mengurangi angka putus sekolah yang tinggi.

4.2 Pola Hubungan Angka Putus Sekolah Usia SMA di Indonesia dengan Faktor-Faktor yang Diduga Berpengaruh

Pola hubungan antara angka putus sekolah usia SMA dengan masing-masing faktor yang diduga mempengaruhinya dapat diketahui dengan *scatterplot*. Pola hubungan ini bertujuan untuk melihat metode yang sesuai untuk memodelkan angka putus sekolah usia SMA di Indonesia. Berdasarkan hasil *scatterplot* dapat diketahui bagaimana pola hubungan yang terbentuk, apakah membentuk suatu pola tertentu atau tidak. Jika *scatterplot* antara angka putus sekolah usia SMA dengan faktor-faktor yang diduga berpengaruh membentuk pola tertentu seperti linier, kuadratik, kubik, atau pola lainnya, maka metode yang digunakan adalah

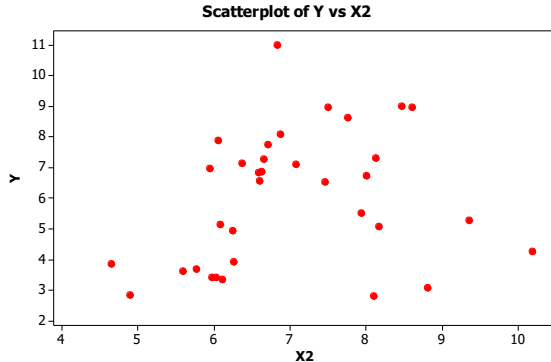
regresi parametrik. Namun, apabila *scatterplot* tidak membentuk pola tertentu maka metode yang digunakan adalah regresi nonparametrik. Apabila *scatterplot* terdapat sebagian membentuk pola tertentu dan sebagian tidak membentuk pola apapun maka metode yang dapat digunakan adalah metode regresi semiparametrik. Berikut ini adalah *scatterplot* antara angka putus sekolah usia SMA dengan faktor-faktor yang diduga mempengaruhinya.



Gambar 4. 2 *Scatterplot* Persentase Penduduk Miskin Dengan Angka Putus Sekolah

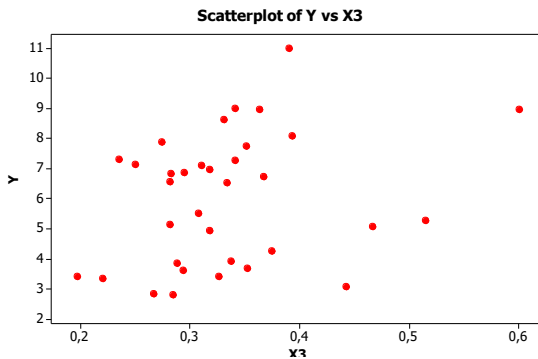
Angka putus sekolah usia SMA di Indonesia dipengaruhi oleh persentase penduduk miskin. Hubungan yang diharapkan antara persentase penduduk miskin dengan angka putus sekolah usia SMA adalah hubungan linier artinya jika persentase penduduk miskin meningkat maka angka putus sekolah usia SMA akan semakin besar. Namun, hal ini bisa saja tidak terjadi karena unit penelitian ini adalah provinsi. Berdasarkan Gambar 4.2 diketahui bahwa antara variabel respon (Y) berupa angka putus sekolah usia SMA dengan variabel prediktor persentase penduduk miskin (X_1) terlihat bahwa tidak membentuk pola tertentu. Oleh sebab itu metode yang dapat digunakan untuk pemodelan angka putus sekolah usia SMA di Indonesia dengan persentase penduduk miskin adalah regresi nonparametrik. Pada Gambar 4.3 berikut

menunjukkan *scatterplot* antara angka putus sekolah usia SMA dengan persentase guru SM/Sederajat terhadap murid (X_2).



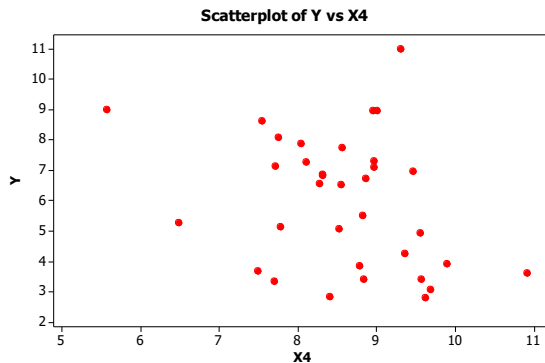
Gambar 4. 3 *Scatterplot* Persentase Guru SM/Sederajat terhadap Murid Dengan Angka Putus Sekolah

Gambar 4.3 menunjukkan pola hubungan angka putus sekolah usia SMA dengan persentase guru SM/Sederajat terhadap murid. Berdasarkan gambar tersebut terlihat bahwa pola data tidak membentuk pola tertentu. Sehingga metode yang digunakan adalah regresi nonparametrik. Pada Gambar 4.4 berikut menunjukkan *scatterplot* antara angka putus sekolah usia SMA dengan persentase jumlah sekolah SM/Sederajat terhadap murid (X_3).



Gambar 4. 4 *Scatterplot* Persentase Jumlah Sekolah SM/Sederajat terhadap Murid Dengan Angka Putus Sekolah

Gambar 4.4 menunjukkan pola hubungan angka putus sekolah usia SMA dengan persentase jumlah sekolah SM/Sederajat terhadap murid. Pada gambar tersebut terlihat pola data tidak membentuk suatu pola tertentu. Sehingga metode yang dapat digunakan adalah metode regresi nonparametrik. Pada Gambar 4.5 berikut menunjukkan *scatterplot* antara angka putus sekolah usia SMA dengan rata-rata lama sekolah penduduk usia 15 tahun ke atas (X_4).



Gambar 4. 5 *Scatterplot* Rata-Rata Lama Sekolah Penduduk Usia 15 Tahun Ke Atas Dengan Angka Putus Sekolah

Gambar 4.5 menunjukkan pola hubungan angka putus sekolah usia SMA dengan persentase jumlah sekolah SM/Sederajat terhadap murid. Pada gambar tersebut terlihat pola data tidak membentuk suatu pola tertentu. Sehingga metode yang dapat digunakan adalah metode regresi nonparametrik.

4.3 Pemilihan Titik Knot Optimum

Berdasarkan pola hubungan yang digambarkan melalui *scatterplot* antara variabel angka putus sekolah usia SMA dengan masing-masing variabel yang diduga mempengaruhinya, dapat diketahui bahwa semua *scatterplot* menunjukkan pola data yang tidak membentuk pola tertentu. Dengan ini, pemodelan angka putus sekolah usia SMA di Indonesia dilakukan menggunakan metode regresi nonparametrik spline. Model regresi nonparametrik spline didapatkan dari titik knot yang optimal.

Untuk pemilihan titik knot yang optimal dilakukan berdasarkan nilai GCV yang paling minimum. Berikut ini merupakan pemilihan titik knot optimal dengan satu knot, dua knot, tiga knot dan kombinasi knot.

4.3.1 Pemilihan Titik Knot dengan Satu Titik Knot

Pemilihan titik knot optimal dimulai dengan pemilihan satu titik knot. Estimasi model regresi nonparametrik splinedengan satu knot pada kasus angka putus sekolah usia SMA di Indonesia adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - k_2)_+^1 + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - k_3)_+^1 + \hat{\beta}_7 x_4 + \hat{\beta}_8 (x_4 - k_4)_+^1$$

Tabel 4.2 berikut merupakan ringkasan nilai GCV model regresi nonparametrik spline dengan satu knot.

Tabel 4. 2Nilai GCV dengan Satu Titik Knot

X₁	X₂	X₃	X₄	GCV
11,80	6,46	0,33	7,93	4,44
12,30	6,57	0,34	8,02	4,34
12,81	6,69	0,35	8,11	4,22
13,31	6,80	0,35	8,20	4,10
13,81	6,91	0,36	8,29	4,07
14,31	7,03	0,37	8,38	4,10
14,82	7,14	0,38	8,47	4,14
15,32	7,25	0,39	8,56	4,18
15,82	7,37	0,39	8,65	4,26
16,33	7,48	0,40	8,75	4,37

Berdasarkan Tabel 4.2 diketahui bahwa nilai GCV terkecil untuk model regresi nonparametrik spline dengan satu knot adalah sebesar 4,07. Satu titik knot pada variabel persentase penduduk miskin (X₁) berada pada titik knot 13,81; variabel persentase guru SM/ sederajat terhadap murid berada pada titik knot 6,91; variabel persentase jumlah sekolah SM/ Sederajat terhadap murid berada pada titik knot 0,36 serta variabel rata-rata lama sekolah penduduk usia 15 tahun keatas berada pada titik knot 8,29.

4.3.2 Pemilihan Titik Knot dengan Dua Titik Knot

Pemilihan titik knot selanjutnya menggunakan dua titik knot. Estimasi model regresi nonparametrik spline dengan dua titik knot pada angka putus sekolah usia SMA di Indonesia adalah sebagai berikut,

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - k_3)_+^1 + \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - k_5)_+^1 + \hat{\beta}_9 (x_3 - k_6)_+^1 + \hat{\beta}_{10} x_4 + \hat{\beta}_{11} (x_4 - k_7)_+^1 + \hat{\beta}_{12} (x_4 - k_8)_+^1$$

Tabel 4.3 merupakan ringkasan nilai GCV untuk model regresi nonparametrik spline dengan dua titik knot. Pada Tabel 4.3 diketahui bahwa nilai GCV terkecil untuk model regresi nonparametrik spline dengan satu knot adalah sebesar 3,86. Satu titik knot pada variabel persentase penduduk miskin (X_1) berada pada titik knot 3,75 dan 13,81; variabel persentase guru SM/ sederajat terhadap murid berada pada titik knot 4,66 dan 6,91; variabel persentase jumlah sekolah SM/ Sederajat terhadap murid berada pada titik knot 0,20 dan 0,36 serta variabel rata-rata lama sekolah penduduk usia 15 tahun keatas berada pada titik knot 6,48 dan 8,29.

Tabel 4.3 Nilai GCV dengan Dua Titik Knot

X_1	X_2	X_3	X_4	GCV
3,75	4,66	0,20	6,48	4,44
11,80	6,46	0,33	7,93	
3,75	4,66	0,20	6,48	4,22
12,81	6,69	0,35	8,11	
3,75	4,66	0,20	6,48	4,10
13,31	6,80	0,35	8,20	
3,75	4,66	0,20	6,48	4,07
13,81	6,91	0,36	8,29	
3,75	4,66	0,20	6,48	4,10
14,31	7,03	0,37	8,38	
3,75	4,66	0,20	6,48	4,14
14,82	7,14	0,38	8,47	
3,75	4,66	0,20	6,48	4,18
15,32	7,25	0,39	8,56	

Tabel 4. 3 Nilai GCV dengan Dua Titik Knot (Lanjutan)

X₁	X₂	X₃	X₄	GCV
3,75	4,66	0,20	6,48	4,26
15,82	7,37	0,39	8,65	
3,75	4,66	0,20	6,48	4,37
16,33	7,48	0,40	8,75	
3,75	4,66	0,20	6,48	4,49
16,83	7,59	0,41	8,80	

4.3.3 Pemilihan Titik Knot dengan Tiga Titik Knot

Estimasi model regresi nonparametrik spline dengan tiga titik knot pada angka putus sekolah usia SMA di Indonesia adalah sebagai berikut.

$$\begin{aligned} \hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_1)_+^1 + \hat{\beta}_3 (x_1 - k_2)_+^1 + \hat{\beta}_4 (x_1 - k_3)_+^1 + \hat{\beta}_5 x_2 + \\ & \hat{\beta}_6 (x_2 - k_4)_+^1 + \hat{\beta}_7 (x_2 - k_5)_+^1 + \hat{\beta}_8 (x_2 - k_6)_+^1 + \hat{\beta}_9 x_3 + \\ & \hat{\beta}_{10} (x_3 - k_7)_+^1 + \hat{\beta}_{11} (x_3 - k_8)_+^1 + \hat{\beta}_{12} (x_3 - k_9)_+^1 + \hat{\beta}_{13} x_4 + \\ & \hat{\beta}_{14} (x_4 - k_{10})_+^1 + \hat{\beta}_{15} (x_4 - k_{11})_+^1 + \hat{\beta}_{16} (x_4 - k_{12})_+^1 \end{aligned}$$

Tabel 4.4 merupakan ringkasan nilai GCV model regresi nonparametrik splinedengan tiga titik knot.

Tabel 4.4Nilai GCV dengan Tiga Titik Knot

X₁	X₂	X₃	X₄	GCV
17,33	7,70	0,42	8,93	3,85
18,84	8,04	0,44	9,20	
20,35	8,38	0,47	9,47	
17,33	7,70	0,42	8,93	2,88
18,84	8,04	0,44	9,20	
20,85	8,49	0,48	9,56	
17,33	7,70	0,42	8,93	2,63
18,84	8,04	0,44	9,20	
21,36	8,61	0,49	9,65	
17,33	7,70	0,42	8,93	2,97
18,84	8,04	0,44	9,20	
21,86	8,72	0,49	9,74	

Tabel 4. 4 Nilai GCV Tiga dengan Titik Knot (Lanjutan)

X₁	X₂	X₃	X₄	GCV
17,33	7,70	0,42	8,93	3,52
18,84	8,04	0,44	9,20	
22,36	8,83	0,50	9,83	
17,33	7,70	0,42	8,93	3,86
18,84	8,04	0,44	9,20	
22,87	8,95	0,51	9,92	
17,33	7,70	0,42	8,93	3,86
18,84	8,04	0,44	9,20	
23,37	9,06	0,52	10,01	
17,33	7,70	0,42	8,93	3,77
18,84	8,04	0,44	9,20	
23,87	9,17	0,53	10,10	

Berdasarkan Tabel 4.4 diketahui bahwa nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan tiga titik knot adalah 2,63. Nilai GCV tersebut diperoleh pada titik knot 17,33; 18,84 dan 21,36 untuk variabel persentase penduduk miskin, titik knot 7,70; 8,04 dan 8,61 untuk variabel persentase guru SM/Sederajat terhadap murid, titik knot 0,42; 0,44 dan 0,49 untuk variabel persentase jumlah sekolah SM/Sederajat terhadap murid, dan titik knot 8,93; 9,20 dan 9,65 untuk variabel rata-rata lama sekolah penduduk usia 15 tahun keatas. Nilai GCV yang dihasilkan dengan menggunakan tiga titik knot lebih kecil daripada nilai GCV yang dihasilkan menggunakan satu titik knot atau dua titik knot.

4.3.4 Pemilihan Titik Knot dengan Kombinasi Titik Knot

Pemilihan titik knot optimal dengan satu titik knot, dua titik knot dan tiga titik knot maka langkah selanjutnya adalah pemilihan titik knot dengan kombinasi titik knot. Pemilihan titik knot optimal pada kombinasil knot dengan mencari nilai GCV yang paling minimum. Berikut ini merupakan nilai GCV dengan kombinasi knot.

Tabel 4.5 Nilai GCV dengan Kombinasi Knot

X₁	X₂	X₃	X₄	GCV
13,81	4,66	0,42	6,48	3,53
	6,91	0,44	8,29	
		0,49		
13,81	4,66	0,42	8,93	2,72
	6,91	0,44	9,20	
		0,49	9,65	
13,81	7,70			4,57
	8,04	0,36	8,29	
	8,61			
13,81	7,70		8,93	2,36
	8,04	0,36	9,20	
	8,61		9,65	

Berdasarkan Tabel 4.5 diketahui bahwa nilai GCV minimum dari kombinasi knot adalah 2,36. Nilai GCV tersebut dihasilkan apabila digunakan kombinasi knot (1,3,1,3). Titik knot yang digunakan adalah 13,81 untuk variabel persentase penduduk miskin, titik knot untuk variabel persentase guru SM/Sederajat terhadap murid adalah 7,70; 8,04; dan 8,61 titik knot untuk variabel persentase jumlah sekolah SM/Sederajat terhadap murid adalah 0,36 serta titik knot untuk variabel rata-rata lama sekolah penduduk usia 15 tahun keatas adalah 8,93; 9,20 dan 9,65. Selanjutnya akan dilakukan pemilihan model terbaik berdasarkan nilai GCV yang paling minimum.

4.3.5 Pemilihan Model Terbaik

Berdasarkan hasil pembahasan sebelumnya telah didapatkan titik knot pada masing-masing variabel dengan menggunakan satu knot, dua knot, tiga knot, dan kombinasi knot. Selanjutnya akan dilakukan pemilihan model terbaik berdasarkan nilai GCV yang paling minimum dari nilai GCV yang telah dihasilkan sebelumnya. Berikut ini adalah hasil nilai GCV yang telah didapatkan pada pembahasan sebelumnya.

Tabel 4.6Perbandingan Nilai GCV dari 1 Titik Knot, 2 Titik Knot, 3 Titik Knot, dan Kombinasi Knot

Banyaknya Titik Knot	GCV
Satu Knot	4,07
Dua Knot	4,07
Tiga Knot	2,63
Kombinasi Knot	2,36

Berdasarkan Tabel 4.6 terlihat bahwa nilai GCV yang paling minimum adalah 2,36 dengan kombinasi titik knot. Maka dapat disimpulkan bahwa model terbaik dapat dibentuk dengan menggunakan kombinasi knot, Kombinasi knot yang akan digunakan adalah kombinasi knot (1,3,1,3).

4.3.6 Penaksiran Parameter Model Regresi Nonparametrik Spline

Pemodelan Angka Putus Sekolah usia SMA dengan regresi nonparametrik spline didapatkan dengan menggunakan titik knot optimal. Pemilihan titik knot optimal yang sudah dilakukan sebelumnya menghasilkan pemodelan terbaik angka putus sekolah usia SMA dan faktor-faktor yang diduga mempengaruhinya dilakukan dengan menggunakan titik knot kombinasi dengan knot masing-masing variabel secara berturut-turut adalah 1,3,1,3. Estimasi parameter regresi nonparametrik spline untuk memodelkan Angka Putus Sekolah usia SMA dan faktor-faktor yang mempengaruhinya adalah sebagai berikut.

$$\begin{aligned}\hat{y} = & -1,71 + 0,01x_1 - 0,01(x_1 - 13,81)_+^1 + 1,43x_2 - \\ & 8,01(x_2 - 7,70)_+^1 + 11,87(x_2 - 8,04)_+^1 - 10,32(x_2 - 8,61)_+^1 + \\ & 7,89x_3 - 13,59(x_3 - 0,36)_+^1 - 0,46x_4 + 26,22(x_4 - 8,93)_+^1 - \\ & 48,65(x_4 - 9,20)_+^1 + 24,94(x_4 - 9,65)_+^1\end{aligned}$$

4.4 Pengujian Parameter Model Regresi Nonparametrik Spline

Pengujian parameter dilakukan untuk mengetahui apakah parameter yang didapatkan signifikan atau tidak pada pemodelan Angka Putus Sekolah usia SMA. Pengujian parameter dilakukan

secara serentak dan individu. Pengujian serentak digunakan untuk mengetahui apakah terdapat parameter yang signifikan atau tidak. Untuk pengujian individu digunakan untuk mengetahui variabel yang memberikan pengaruh signifikan.

4.4.1 Uji Serentak

Uji serentak digunakan untuk mengetahui apakah parameter model regresi signifikan atau tidak. Berikut merupakan hipotesis dari pengujian parameter serentak.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{12} = 0$$

$$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, 12$$

Statistik uji menggunakan uji F. Berikut merupakan hasil uji serentak.

Tabel 4.7 Analisis Ragam Uji Serentak

Sumber	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F hitung	P-value
Regresi	12	123,58	10,30	7,06	$5,71 \times 10^{-5}$
Error	21	30,63	1,46		
Total	33	154,21			

Berdasarkan Tabel 4.7 dapat diketahui bahwa nilai *p-value* adalah sebesar $5,71 \times 10^{-5}$ dan nilai ini kurang dari nilai α (0,05) sehingga dapat disimpulkan bahwa tolak H_0 . Hal ini menunjukkan bahwa minimal terdapat satu parameter yang tidak sama dengan nol. Setelah didapatkan hasil yang signifikan pada pengujian parameter model secara serentak, maka langkah selanjutnya adalah melakukan pengujian parameter model secara individu.

4.4.2 Uji Individu

Pengujian signifikansi parameter model secara individu dilakukan untuk mengetahui parameter mana saja yang berpengaruh signifikan. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 12$$

Berikut merupakan hasil pengujian secara parsial pada masing-masing parameter setiap prediktor.

Tabel 4.8Hasil Uji Individu

Variabel	Parameter	Estimasi	t	p-value	Ket,
X ₁	β_0	-1,71	-0,32	0,74869	Tidak Signifikan
	β_1	0,01	0,08	0,93652	Tidak Signifikan
	β_2	-0,01	-0,11	0,91532	Tidak Signifikan
X ₂	β_4	1,43	3,45	0,00238	Signifikan
	β_5	-8,01	-2,51	0,02047	Signifikan
	β_6	11,87	2,46	0,02268	Signifikan
	β_7	-10,32	-2,84	0,00978	Signifikan
X ₃	β_8	7,89	1,32	0,20195	Tidak Signifikan
	β_9	-13,59	-1,24	0,22988	Tidak Signifikan
X ₄	β_{10}	-0,46	-0,83	0,41763	Tidak Signifikan
	β_{11}	26,22	4,24	0,00036	Signifikan
	β_{12}	-48,65	-4,93	$7,12 \times 10^{-5}$	Signifikan
	β_{13}	24,94	5,11	$4,57 \times 10^{-5}$	Signifikan

Apabila nilai *p-value* kurang dari α yaitu 0,05 maka parameter signifikan, Berdasarkan Tabel 4.8 diketahui bahwa dengan tingkat signifikansi 5%, terdapat dua variabel yang berpengaruh signifikan terhadap model yaitu persentase guru SM/Sederajat terhadap murid (X₂), dan rata-rata lama sekolah penduduk usia 15 tahun ke atas (X₄). Namun, terdapat dua variabel yang tidak berpengaruh signifikan terhadap model yaitu persentase penduduk miskin (X₁) dan persentase jumlah sekolah SM/Sederajat (X₃).

4.5 Pemodelan Rergresi Nonparametrik SplineOptimal dengan Dua Variabel Prediktor

Pada pengujian signifikan parameter diketahui bahwa hanya dua variabel yang berpengaruh secara signifikan terhadap model

yaitu X_2 dan X_4 . Dengan ini akan dilakukan pemodelan kembali menggunakan dua variabel tersebut. Pemodelan regresi nonparametrik spline juga akan menggunakan satu titik knot, dua titik knot, tiga titik knot dan kombinasi titik knot.

4.5.1 Pemilihan Titik Knot dengan Satu Titik Knot

Estimasi model regresi nonparametrik spline dengan dua variabel prediktor menggunakan satu titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_2 + \hat{\beta}_2 (x_2 - k_1)_+^1 + \hat{\beta}_3 x_4 + \hat{\beta}_4 (x_4 - k_2)_+^1$$

Tabel 4.9 berikut merupakan ringkasan nilai GCV untuk model regresi nonparametrik spline dengan satu knot.

Tabel 4.9 Nilai GCV dengan Satu Titik Knot Pada Dua Variabel Prediktor

X_2	X_4	GCV
6,46	7,93	4,09
6,57	8,02	3,93
6,69	8,11	3,78
6,80	8,20	3,62
6,91	8,29	3,56
7,03	8,38	3,57
7,14	8,47	3,60
7,25	8,56	3,64
7,37	8,65	3,70
7,48	8,75	3,79

Berdasarkan Tabel 4.9 diketahui bahwa nilai GCV terkecil untuk model regresi nonparametrik spline dengan satu knot adalah sebesar 3,56. Satu titik knot pada variabel persentase guru SM/ sederajat terhadap murid berada pada titik knot 6,91 dan variabel rata-rata lama sekolah penduduk usia 15 tahun keatas berada pada titik knot 8,29.

4.5.2 Pemilihan Titik Knot dengan Dua Titik Knot

Berikut ini merupakan model regresi nonparametrik spline dengan dua variabel prediktor menggunakan satu titik knot.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_2 + \hat{\beta}_2 (x_2 - k_1)_+^1 + \hat{\beta}_3 (x_2 - k_2)_+^1 + \hat{\beta}_4 x_4 + \hat{\beta}_5 (x_4 - k_3)_+^1 + \hat{\beta}_6 (x_4 - k_4)_+^1$$

Tabel 4.10 berikut merupakan ringkasan nilai GCV untuk model regresi nonparametrik spline dengan dua knot.

Tabel 4.10 Nilai GCV dengan Dua Titik Knot Pada Dua Variabel

X₂	X₄	GCV
8,27	9,38	4,33
9,85	10,65	
8,27	9,38	4,33
9,96	10,74	
8,27	9,38	4,33
10,07	10,83	
8,27	9,38	4,45
10,19	10,92	
8,38	9,47	3,66
8,49	9,56	
8,38	9,47	3,51
8,61	9,65	
8,38	9,47	3,88
8,72	9,74	
8,38	9,47	4,24
8,83	9,83	
8,38	9,47	4,42
8,95	9,92	
8,38	9,47	4,42
9,06	10,01	

Pada Tabel 4.10 diketahui bahwa nilai GCV terkecil untuk model regresi nonparametrik spline dengan dua knot adalah sebesar 3,51. Nilai ini diperoleh dari titik knot optimal pada setiap variabel. Dua titik knot pada variabel persentase guru SM/ sederajat terhadap murid berada pada titik knot 8,38 dan 8,61;

serta variabel rata-rata lama sekolah penduduk usia 15 tahun keatas berada pada titik knot 9,47 dan 9,65.

4.5.3 Pemilihan Titik Knot dengan Tiga Titik Knot

Estimasi model Angka Putus Sekolah usia SMA di Indonesia dengan dua variabel prediktor menggunakan regresi nonparametrik spline dengan tiga titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_2 + \hat{\beta}_2 (x_2 - k_1)_+^1 + \hat{\beta}_3 (x_2 - k_2)_+^1 + \hat{\beta}_4 (x_2 - k_3)_+^1 + \hat{\beta}_5 x_4 + \hat{\beta}_6 (x_4 - k_4)_+^1 + \hat{\beta}_7 (x_4 - k_5)_+^1 + \hat{\beta}_8 (x_4 - k_6)_+^1$$

Tabel 4.11 berikut merupakan ringkasan nilai GCV untuk model regresi nonparametrik spline dengan tiga knot.

Tabel 4.11Nilai GCV dengan Tiga Titik Knot Pada Dua Variabel Prediktor

X₂	X₄	GCV
7,70	8,93	3,31
8,04	9,20	
10,07	10,83	
7,70	8,93	2,69
8,16	9,29	
8,27	9,38	
7,70	8,93	2,47
8,16	9,29	
8,38	9,47	
7,70	8,93	1,95
8,16	9,29	
8,49	9,56	
7,70	8,93	1,78
8,16	9,29	
8,61	9,65	
7,70	8,93	2,19
8,16	9,29	
8,72	9,74	
7,70	8,93	3,26
8,16	9,29	
8,95	9,92	

Berdasarkan Tabel 4.11 diketahui bahwa nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan tiga titik knot adalah 1,78. Nilai GCV tersebut diperoleh pada titik knot 7,70; 8,16 dan 8,61 untuk variabel persentase guru SM/Sederajat terhadap murid dan titik knot 8,93; 9,29 dan 9,65 untuk variabel rata-rata lama sekolah penduduk usia 15 tahun keatas.

4.5.4 Pemilihan Titik Knot dengan Kombinasi Knot

Berikut ini merupakan ringkasan nilai GCV diantara GCV yang paling minimum menggunakan kombinasi knot disajikan dalam Tabel 4.12 sebagai berikut.

Tabel 4. 12Nilai GCV dengan Kombinasi Knot Pada Dua Variabel Prediktor

X₂	X₄	GCV
6,91	9,47 9,65	2,50
8,38 8,61	8,29	4,76
8,38 8,61	9,47 9,65	3,51
8,38 8,61	8,93 9,29 9,65	2,38
7,70 8,16 8,61	8,29	3,77
7,70 8,16 8,61	9,47 9,65	2,47

Berdasarkan Tabel 4.12 diketahui bahwa nilai GCV minimum dari kombinasi knot adalah 2,38. Nilai GCV tersebut dihasilkan apabila digunakan kombinasi knot (2,3).Titik knot

untuk variabel persentase guru SM/Sederajat terhadap murid adalah 8,38 dan 8,61, serta titik knot untuk variabel rata-rata lama sekolah penduduk usia 15 tahun keatas adalah 8,93; 9,29 dan 9,65.

4.5.5 Pemilihan Model Terbaik

Pemilihan model terbaik berdasarkan nilai GCV yang paling minimum dari nilai GCV yang telah dihasilkan sebelumnya, Berikut ini adalah hasil nilai GCV yang telah didapatkan pada pembahasan sebelumnya,

Tabel 4. 13Nilai GCV Minimum Dari Hasil Titik Knot Optimal Pada Dua Variabel Prediktor

Banyaknya Titik Knot	GCV
Satu Knot	3,56
Dua Knot	3,51
Tiga Knot	1,78
Kombinasi Knot	2,38

Berdasarkan Tabel 4.13 terlihat bahwa nilai GCV yang paling minimum adalah 1,78 dengan banyaknya titik knot adalah tiga knot. Maka dapat disimpulkan bahwa model terbaik dapat dibentuk dengan menggunakan tiga titik knot.

4.5.6 Penaksiran Parameter Model Regresi Nonparametrik Spilne

Setelah mendapatkan model terbaik yaitu dengan tiga titik knot maka akan dilakukan estimasi parameter model. Berikut ini adalah model regresi nonparametrik spline terbaik.

$$\begin{aligned}\hat{y} = & -1,48 + 1,71x_2 - 7,41(x_2 - 7,70)_+^1 + 12(x_2 - 8,16)_+^1 - \\ & 11,37(x_2 - 8,61)_+^1 - 0,41x_4 + 15,60(x_4 - 8,93)_+^1 - \\ & 39,62(x_4 - 9,29)_+^1 + 26,57(x_4 - 9,65)_+^1\end{aligned}$$

4.6 Pengujian Parameter Model Regresi Nonparametrik Spline

Pengujian parameter dilakukan untuk mengetahui apakah estimasi parameter yang didapatkan berpengaruh signifikan atau tidak pada pemodelan Angka Putus Sekolah usia SMA. Pengujian

dilakukan secara serentak dan individu. Apabila hasil uji signifikansi secara serentak menunjukkan bahwa terdapat minimal satu parameter yang signifikan, maka pengujian dilanjutkan secara individu.

4.6.1 Uji Serentak

Ujiserentak digunakan untuk mengetahui apakah parameter model regresi signifikan atau tidak. Berikut merupakan hipotesis dari pengujian parameter serentak.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_8 = 0$$

$$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, 8$$

Statistik uji menggunakan uji F. Berikut merupakan hasil uji serentak.

Tabel 4. 14 Analisis Ragam Uji Serentak Dua Variabel Prediktor

Sumber	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	Fhitung	P-value
Regresi	8	121,42	15,77	11,57363	$1,03 \times 10^{-6}$
Error	25	32,78	1,31		
Total	33	154,21			

Berdasarkan Tabel 4.14 dapat diketahui bahwa nilai *p-value* adalah sebesar $1,03 \times 10^{-6}$ dan nilai ini kurang dari nilai α (0,05) sehingga dapat disimpulkan bahwa tolak H_0 . Hal ini menunjukkan bahwa minimal terdapat satu parameter yang tidak sama dengan nol

4.6.2 Uji Individu

Pengujian signifikansi parameter model secara individu dilakukan untuk mengetahui parameter mana saja yang berpengaruh signifikan. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 8.$$

Berikut merupakan hasil pengujian secara parsial pada masing-masing parameter setiap prediktor.

Tabel 4. 15Hasil Uji Individu

Variabel	Parameter	Estimasi	t- hitung	P-value	Ket
Constant	β_0	-1,48	0,36	0,72505	Tidak Signifikan
X ₂	β_1	1,71	5,02	$0,73 \times 10^{-6}$	Signifikan
	β_2	-7,41	-3,31	0,00280	Signifikan
	β_3	12	3,14	0,00428	Signifikan
	β_4	-11,37	-3,86	0,00071	Signifikan
X ₄	β_5	-0,41	-0,95	0,35049	Tidak Signifikan
	β_6	15,60	4,42	0,00016	Signifikan
	β_7	-39,62	-5,69	$6,3 \times 10^{-6}$	Signifikan
	β_8	26,57	5,95	$3,3 \times 10^{-6}$	Signifikan

Berdasarkan Tabel 4.15 diketahui bahwa kedua variabel prediktor yang digunakan yaitu persentase guru SM/Sederajat terhadap murid (X₂) dan rata-rata lama sekolah penduduk usia 15 keatas (X₄) berpengaruh signifikan terhadap Angka Putus Sekolah usia SMA di Indonesia.

4.7 Pengujian Asumsi Residual

Penggunaan metode regresi nonparametrik spline harus memenuhi asumsi residual IIDN. Hal ini berarti bahwa residual dari model regresi nonparametrik yang telah diperoleh harus memenuhi asumsi identik, independen dan berdistribusi normal. Oleh karena itu, akan dilakukan tiga pemeriksaan terhadap residual model regresi nonparametrik splinesebagai berikut.

4.7.1 Uji Asumsi Identik

Salah satu asumsi regresi yang harus dipenuhi adalah homogenitas variansi dari residual (homoskedastisitas). Homoskedastisitas berarti bahwa variansi residual bersifat konstan (tetap) atau disebut identik. Salah satu cara untuk mendeteksi adanya heteroskedastisitas adalah dengan menggunakan uji *Glejser*. Berikut merupakan hipotesis yang digunakan dalam uji Glejser.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{34}^2 = \sigma^2$$

$$H_1 : \text{paling sedikit ada satu } \sigma_i^2 \neq \sigma^2$$

Berikut merupakan hasil uji Glejser pada residual.

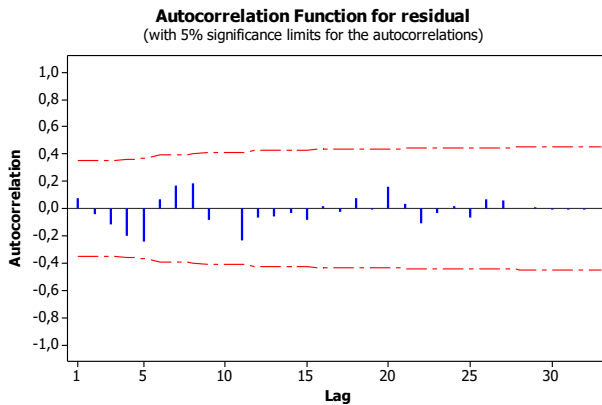
Tabel 4. 16 Analisis Ragam Uji Glejser

Sumber	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F hitung	P-value
Regresi	8	2,57	0,32	0,8606	0,561129
Error	25	9,34	0,37		
Total	33	11,91			

Berdasarkan Tabel 4.16 dapat diketahui bahwa nilai *p-value* sebesar 0,561129 bernilai lebih besar dibandingkan dengan nilai α yaitu 0,05. Hal ini menunjukkan H_0 gagal tolak dan dapat disimpulkan bahwa tidak terjadi kasus heteroskedastisitas.

4.7.2 Pemeriksaan Asumsi Independen

Asumsi berikutnya yang harus terpenuhi adalah residual yang independen. Residual yang independen adalah tidak terjadi autokorelasi antar residual. Salah satu cara untuk mengetahui apakah terjadi kasus autokorelasi atau tidak adalah dengan melihat *plot* dari *Autocorrelation Function* (ACF). Berikut ini adalah *plot* dari *Autocorrelation Function* (ACF).



Gambar 4.6 Plot ACF Residual

Apabila terdapat autokorelasi yang melewati atau keluar dari interval konfidensi maka dapat dinyatakan terjadi autokorelasi. Pada Gambar 4.6 terlihat bahwa tidak terdapat autokorelasi yang melewati atau keluar dari interval konfidensi. Sehingga dapat disimpulkan bahwa tidak terjadi kasus autokorelasi.

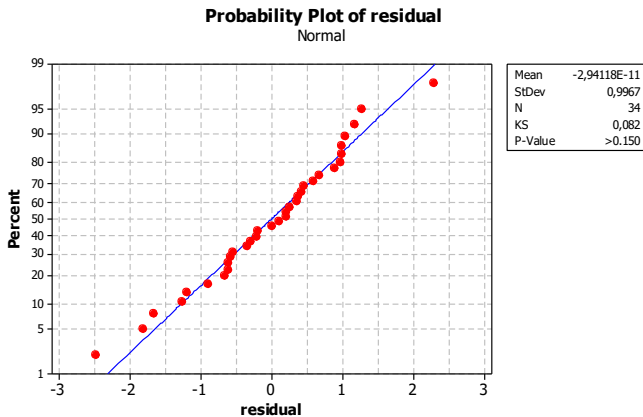
4.7.3 Uji Asumsi Distribusi Normal

Asumsi terakhir yang harus dipenuhi adalah residual yang dihasilkan berdistribusi normal, Pengujian apakah residual yang dihasilkan berdistribusi normal atau tidak dengan menggunakan uji *Kolmogorov-Smirnov*. Berikut merupakan hipotesis yang digunakan dalam uji Kolmogorov-Smirnov.

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Berikut merupakan *Normal Probability Plot* yang dihasilkan.



Gambar 4.7 Hasil Uji *Kolmogorov-Smirnov*

Berdasarkan Gambar 4.7 diketahui bahwa nilai *Kolmogorov-Smirnov* adalah 0,082. Nilai tersebut lebih kecil jika dibandingkan dengan nilai $q_{(1-\alpha)}$ yaitu 0,227. Dengan membandingkan nilai tersebut maka H_0 gagal ditolak dan disimpulkan bahwa residual berdistribusi normal. Selain itu, nilai *p-value* lebih dari 0,150

bernilai lebih besar dibandingkan dengan nilai α (0,05). Perbandingan ini menghasilkan kesimpulan yang sama bahwa residual dari model regresi nonparametrik spline berdistribusi normal.

4.8 Nilai Koefisien Determinasi (R^2)

Koefisien determinasi menggambarkan bagian dari variasi total yang didapat diterangkan model. Berikut adalah perhitungan koefisien determinansi berdasarkan hasil Tabel 4.17.

$$\begin{aligned} R^2 &= \frac{SS_{regresi}}{SS_{total}} \times 100\% \\ &= \frac{121,42}{154,21} \times 100\% \\ &= 78,74\% \end{aligned}$$

Hasil perhitungan koefisien determinasi menunjukkan nilai 78,74%. Nilai ini menunjukkan bahwa model regresi nonparametrik spline yang terbentuk mampu menjelaskan keragaman Angka Putus Sekolah usia SMA di Indonesia sebesar 78,74%, sedangkan sisanya dijelasnya oleh variabel lain.

4.9 Interpretasi Model Regresi Nonparametrik Spline

Model regresi nonparametrik splineterbaik adalah dengan tiga titik knot. Berikut ini merupakan persamaan model yang diperoleh dari tiga titik knot.

$$\begin{aligned} \hat{y} &= -1,48 + 1,71x_2 - 7,41(x_2 - 7,70)_+^1 + 12(x_2 - 8,16)_+^1 - \\ &\quad 11,37(x_2 - 8,61)_+^1 - 0,41x_4 + 15,60(x_4 - 8,93)_+^1 - \\ &\quad 39,62(x_4 - 9,29)_+^1 + 26,57(x_4 - 9,65)_+^1 \end{aligned}$$

Model tersebut memiliki dua variabel dimana kedua variabel tersebut berpengaruh signifikan terhadap angka putus sekolah usia SMA. Interpretasi model bertujuan untuk mengetahui seberapa besar pengaruh masing-masing variabel terhadap angka putus sekolah usia SMA. Interpretasi dari model tersebut adalah sebagai berikut.

1. Jika variabel X_2 dianggap konstan maka pengaruh persentase guru SM/Sederajat terhadap murid (X_2) kepada angka putus sekolah usia SMA di Indonesia sebagai berikut.

$$\hat{y} = -1,48 + 1,71X_2 - 7,41(X_2 - 7,70) -$$

$$12(X_2 - 8,16) - 11,37(X_2 - 8,61)$$

$$= \begin{cases} -1,48 + 1,71X_2 & , \quad X_2 < 7,70 \\ 55,64 - 5,71X_2 & , \quad 7,70 \leq X_2 < 8,16 \\ -42,20 + 6,29X_2 & , \quad 8,16 \leq X_2 < 8,61 \\ 55,63 - 5,08X_2 & , \quad X_2 > 8,61 \end{cases}$$

Pada model diatas dapat diinterpretasikan bahwa apabila persentase guru SM/Sederajat kurang dari 7,70% dan persentase guru SM/Sederajat mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan meningkat sebesar 1,71%. Wilayah diantaranya adalah provinsi Sumatera Utara, Riau, Jambi, Sumatera Selatan, Lampung, Kep Bangka Belitung, Kep Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Bali, Kalimantan Barat, Kalimantan Selatan, Kalimantan Timur, Kalimantan, Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Gorontalo, Sulawesi Barat, dan Papua Barat.

Apabila persentase guru SM/Sederajat terhadap murid diantara 7,70% sampai dengan 8,16% dan apabila persentase guru SM/Sederajat terhadap murid mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan menurun sebesar 5,71%. Wilayah diantaranya adalah provinsi Sumatera Barat, Bengkulu, DI Yogyakarta, Nusa Tenggara Timur, Sulawesi Tenggara.

Apabila persentase guru SM/Sederajat terhadap murid diantara 8,16% sampai dengan 8,61% dan persentase guru SM/Sederajat mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan meningkat sebesar 6,29%. Wilayah diantaranya adalah provinsi Nusa Tenggara Barat, Kalimantan Tengah.

Apabila persentase guru SM/Sederajat terhadap murid lebih dari 8,61% dan apabila persentase guru SM/Sederajat terhadap murid mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan menurun sebesar 5,08%, Wilayah diantaranya adalah provinsi Aceh, Maluku, Maluku Utara, dan Papua.

Berbasarkan model yang diperoleh, maka didapatkan peta persebaran angka putus sekolah usia SMA berdasarkan persentase guru SM/Sederajat terhadap murid seperti pada Gambar 4,8 berikut.



Gambar 4.8Peta Persebaran Angka Putus Sekolah Usia SMA Berdasarkan Persentase Guru SM/Sederajat Terhadap Murid

2. Jika variabel X_4 dianggap konstan maka pengaruh rata-rata lama sekolah penduduk usia 15 ke atas (X_4) kepada angka putus sekolah usia SMA di Indonesia sebagai berikut,

$$\hat{y} = -1,48 - 0,41X_4 + 15,60(X_4 - 8,93) - 39,62(X_4 - 9,29) + 26,57(X_4 - 9,65)$$

$$= \begin{cases} -1,48 - 0,41X_4 & , X_4 < 8,93 \\ -140,75 + 15,19X_4 & , 8,93 \leq X_4 < 9,29 \\ 227,27 - 24,43X_4 & , 9,29 \leq X_4 < 9,65 \\ 29,19 + 2,15X_4 & , X_4 \geq 9,65 \end{cases}$$

Pada model diatas dapat diinterpretasikan bahwa apabila rata-rata lama sekolah penduduk usia 15 tahun ke atas diantara 8,93%

sampai dengan 9,29% dan rata-rata lama sekolah penduduk usia 15 tahun ke atas mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan meningkat sebesar 15,19%. Wilayah diantaranya adalah provinsi Sumatera Barat, Riau, Kalimantan Utara, dan Maluku Utara.

Apabila rata-rata lama sekolah penduduk usia 15 tahun ke atas diantara 9,29% sampai dengan 9,65% dan apabila persentase guru SM/Sederajat mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan menurun sebesar 24,43%. Wilayah diantaranya adalah provinsi Aceh, Sumatera Utara, DI Yogyakarta, Kalimantan Timur, Sulawesi Utara dan Papua Barat.

Apabila rata-rata lama sekolah penduduk usia 15 tahun ke atas lebih dari 9,65%, dan rata-rata lama sekolah penduduk usia 15 tahun ke atas mengalami kenaikan sebesar 1% maka angka putus sekolah usia SMA akan meningkat sebesar 2,15%, Wilayah diantaranya adalah provinsi Kep Riau, DKI Jakarta, dan Maluku.

Berbasarkan model yang diperoleh, maka didapatkan peta persebaran angka putus sekolah usia SMA berdasarkan persentase guru SM/Sederajat terhadap murid seperti pada Gambar 4.9 berikut.



Gambar 4, 9Peta Persebaran Angka Putus Sekolah Usia SMA Berdasarkan Rata-Rata Lama Sekolah Penduduk Usia 15 Ke atas

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut,

1. Rata-rata Angka Putus Sekolah usia SMA di Indonesia adalah sebesar 5,98% dengan varian sebesar 4,67. Nilai angka putus sekolah usia SMA terendah berada pada provinsi DI Yogyakarta yaitu sebesar 2,8%. Sementara nilai paling tinggi berada pada provinsi Sulawesi Utara yaitu mencapai 10,97 %. Provinsi yang memiliki persentase angka putus sekolah lebih tinggi dari rata-rata adalah provinsi Maluku Utara, Sulawesi Barat, Gorontalo, Sulawesi Tenggara, Sulawesi Selatan, Sulawesi Tengah, Sulawesi Utara, Kalimantan Utara, Kalimantan Selatan, Nusa Tenggara Timur, Nusa Tenggara Barat, Kep Bangka Belitung, Lampung, Sumatera Selatan, Jambi, Riau, Sumatera Barat, Sumatera Utara dan Aceh.
2. Model regresi nonparametrik splineterbaik untuk pemodelan Angka Putus Sekolah usia SMA di Indonesia didapatkan dengan GCV tiga titik knot. Berikut ini merupakan model regresi nonparametrik spline terbaik untuk angka putus sekolah usia SMA di Indonesia

$$\begin{aligned}\hat{y} = & -1,48 + 1,71x_2 - 7,41(x_2 - 7,70)_+^1 + 12(x_2 - 8,16)_+^1 - \\ & 11,37(x_2 - 8,61)_+^1 - 0,41x_4 + 15,60(x_4 - 8,93)_+^1 - \\ & 39,62(x_4 - 9,29)_+^1 + 26,57(x_4 - 9,65)_+^1\end{aligned}$$

Model ini memiliki nilai GCV minimum yaitu 1,78 dengan koefisien determinasi (R^2) sebesar 78,74%. Variabel yang berpengaruh secara signifikan adalah persentase guru SM/Sederajat terhadap murid dan rata-rata lama sekolah penduduk usia 15 tahun ke atas.

5.2 Saran

Saran yang dapat diberikan pada penelitian selanjutnya dan untuk pemerintah. Untuk peneliti selanjutnya sebaiknya menambahkan faktor-faktor lain yang mempengaruhi angka putus sekolah usia SMA di Indonesia. Saran untuk pemerintah sebaiknya, pemerintah meningkatkan akses dan layanan kualitas pendidikan agar dapat menurunkan angka putus sekolah di Indonesia.

DAFTAR PUSTAKA

- Badan Pusat Statistik (BPS). (2010). *Hasil Sensus Penduduk 2010 Data Agregat per Provinsi*. Jakarta : Badan Pusat Statistik.
- Badan Pusat Statistik (BPS). (2015). *Profil Anak Indonesia 2015*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik (BPS). (2016). *Potret Pendidikan Indonesia Statistik Pendidikan 2016*. Jakarta: Badan Pusat Statistik.
- Badan Pusat Statistik (BPS). (2017). *Angka Putus Sekolah*. Diakses 23 Januari 2017, dari <https://sirusa.bps.go.id/>
- Beeby, C.E. (1982). *Pendidikan Di Indonesia Penilaian dan Pedoman Perencanaan*. Jakarta: LP3ES.
- Budiantara, I.N. (2007). Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik. *Jurnal Matematika, Ilmu Pengetahuan Alam dan Pengajarannya Universitas Negeri Malang*, 36(1), 1-16.
- Budiantara, I.N. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.
- Citrasari, T. (2011). *Pemodelan Angka Putus Sekolah Bagi Anak Usia Wajib Belajar Di Jawa Timur dengan Pendekatan Generalized Poisson Regression*. Tugas Akhir Sarjana, Institut Teknologi Sepuluh Nopember, Surabaya.
- Daniel, W. (1989). *Statistika Nonparametrik Terapan*. Diterjemahkan oleh: Alex Tri Kantjono W. Jakarta: PT Gramedia.
- Deputi Evaluasi Kerja Pembangunan. (2009). *Evaluasi Pelaksanaan Program Wajib Belajar Pendidikan Dasar 9 tahun*. Jakarta: Badan Perencanaan Pembangunan Nasional.
- Draper, N., & Smith, H. (1992). *Applied Regression Analysis, Second Edition*. New York: John Willey & Sons, Inc.
- Eubank, R. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker, Inc.

- Gujarati, N. D. (2003). *Basic Econometrics fourth edition*. New York: The McGraw-Hill Companies, Inc.
- Kementerian Pendidikan dan Kebudayaan (Kemendikbud). (2015). *Rencana Strategis Kementerian Pendidikan dan Kebudayaan 2015-2019*. Jakarta: Kemendikbud.
- Mubarokah, L. (2016). *Pemodelan Angka Putus Sekolah Usia SMP Menggunakan Metode Regresi Nonparametrik Spline di Papua*. Tugas Akhir Sarjana, Institut Teknologi Sepuluh Nopember, Surabaya.
- Musfiquon, M. (2009). *Memberlajarkan yang putus sekolah*. Diakses 27 Februari 2017, dari <http://fiqon-umsida.blogspot.co.id/>.
- OECD. (2015). *Education at a Glance 2015 : OECD Indicators*. Paris: OECD Publishing.
- Pradipta, M. (2013). *Pemodelan Angka Putus Sekolah Usia SMA Di Jawa Timur Dengan Pendekatan Regresi Spline Multivariabel*. Tugas Akhir Sarjana, Institut Teknologi Sepuluh Nopember, Surabaya.
- Farah, M. (2014). *Faktor Penyebab Putus Sekolah dan Dampak Negatifnya Bagi Anak (Studi Kasus di Desa Kalisoro Kecamatan Tawangmangu Kabupaten Karanganyar)*. Tugas Akhir Sarjana, Universitas Muhammadiyah Surakarta, Surakarta.
- Septiana, L. (2011). *Pemodelan Remaja Putus Sekolah Usia SMA di Provinsi Jawa Timur Dengan Menggunakan Metode Regresi Spasial*. Tugas Akhir Sarjana, Institut Teknologi Sepuluh Nopember, Surabaya.
- Walpole, R.E. (1995). *Pengantar Metode Statistika*. Edisi Ketiga, Alih Bahasa : Bambang Sumantri. Jakarta: Penerbit PT Gramedia Pusaka Utama.
- Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods Second Edition*. New York: Addison Wesley.

LAMPIRAN

Lampiran 1. Data Angka Putus Sekolah Usia SMA di Indonesia
dengan Faktor-Faktor yang Mempengaruhi Tahun
2016

Provinsi	Y	X ₁	X ₂	X ₃	X ₄
Aceh	4,24	16,43	10,19	0,37	9,36
Sumatera Utara	6,95	10,27	5,95	0,32	9,46
Sumatera Barat	7,29	7,14	8,13	0,24	8,97
Riau	7,08	7,67	7,08	0,31	8,97
Jambi	6,51	8,37	7,47	0,33	8,55
Sumatera Selatan	6,85	13,39	6,63	0,30	8,32
Bengkulu	5,49	17,03	7,95	0,31	8,82
Lampung	7,25	13,86	6,66	0,34	8,1
Kep, Bangka Belitung	7,87	5,04	6,05	0,27	8,04
Kep, Riau	3,92	5,84	6,26	0,34	9,9
Dki Jakarta	3,59	3,75	5,59	0,29	10,92
Jawa Barat	2,83	8,77	4,90	0,27	8,41
Jawa Tengah	3,34	13,19	6,12	0,22	7,7
Di Yogyakarta	2,8	13,1	8,11	0,28	9,62
Jawa Timur	5,12	11,85	6,08	0,28	7,78
Banten	3,85	5,36	4,66	0,29	8,79
Bali	3,38	4,15	6,03	0,20	8,84
Nusa Tenggara Barat	8,97	16,02	8,47	0,34	7,57
Nusa Tenggara Timur	8,6	22,01	7,77	0,33	7,54
Kalimantan Barat	3,65	8	5,77	0,35	7,49
Kalimantan Tengah	5,06	5,36	8,17	0,47	8,52
Kalimantan Selatan	6,53	4,52	6,61	0,28	8,28
Kalimantan Timur	4,93	6	6,26	0,32	9,55
Kalimantan Utara	8,96	6,99	7,51	0,36	9,01
Sulawesi Utara	10,97	8,2	6,84	0,39	9,31
Sulawesi Tengah	7,74	14,09	6,71	0,35	8,56
Sulawesi Selatan	6,83	9,24	6,59	0,28	8,31
Sulawesi Tenggara	6,72	12,77	8,00	0,37	8,86
Gorontalo	7,13	17,63	6,37	0,25	7,71
Sulawesi Barat	8,08	11,19	6,88	0,39	7,76
Maluku	3,07	19,26	8,81	0,44	9,69

Provinsi	Y	X₁	X₂	X₃	X₄
Maluku Utara	8,96	6,41	8,62	0,60	8,96
Papua Barat	3,41	24,88	5,97	0,33	9,57
Papua	5,25	28,4	9,36	0,51	6,48

Keterangan:

Y = Angka Putus Sekolah Usia SMA

X₁ = Persentase Penduduk Miskin

X₂ = Persentase Guru SM/Sederajat Terhadap Murid

X₃ = Persentase Sekolah SM/Sederajat Terhadap Murid

X₄ = Rata-Rata Lama Sekolah Penduduk Usia 15+

Lampiran 2. Program Pemilihan Titik Knot Optimal dengan Satu Titik Knot Menggunakan Software R

```
GCV1=function(para)
{
  data=read.table("d://data.txt",header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[,2])
  m=ncol(data)-para-1
  dataA=data[, (para+2):q]
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot1=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot1[j,i]=a[j]
    }
  }
  a1=length(knot1[,1])
  knot1=knot1[2:(a1-1),]
  aa=rep(1,p)
  data1=matrix(ncol=m,nrow=p)
  data2=data[,2:q]
  a2=nrow(knot1)
  GCV=rep(NA,a2)
  Rsq=rep(NA,a2)
  for (i in 1:a2)
  {
    for (j in 1:m)
    {
      for (k in 1:p)
      {
        if (data[k,(j+para+1)]<knot1[i,j]) data1[k,j]=0 else
          data1[k,j]=data[k,(j+para+1)]-knot1[i,j]
      }
    }
  }
}
```

```

}
}
mx=cbind(aa,data2,data1)
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in (1:p))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsq[i]=(SSR/(SSE+SSR))*100
MSE=SSE/p
A=mx%*%C%*%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsq=as.matrix(Rsq)
cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsq dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (Rsq)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
print(max(Rsq))

```

```
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 1 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d://validasi_GCV1.csv")
write.csv(Rsq, file="d://validasi_Rsq1.csv")
write.csv(knot1, file="d://validasi_knot1.csv")
}
```

Lampiran 3. Program Pemilihan Titik Knot Optimal dengan Dua Titik Knot Menggunakan *Software R*

```
GCV2=function(para)
{
  data=read.table("d://data.txt", header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-1
  F=matrix(0,nrow=p,ncol=p)
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(data[, (i+1)]),max(data[, (i+1)]),length.out=50)
      knot[j,i]=a[j]
    }
  }
  z=(nk*(nk-1)/2)
  knot2=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot1=rbind(rep(NA,2))
    for ( j in 1:(nk-1))
    {
      for (k in (j+1):nk)
      {
        xx=cbind(knot[j,i],knot[k,i])
        knot1=rbind(knot1,xx)
      }
    }
    knot2=cbind(knot2,knot1)
  }
  knot2=knot2[2:(z+1),2:(2*m+1)]
  aa=rep(1,p)
```

```

data2=matrix(ncol=(2*m),nrow=p)
data1=data[,2:q]
a1=length(knot2[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:(2*m))
  {
    if (mod(j,2)==1) b=floor(j/2)+1 else b=j/2
    for (k in 1:p)
    {
      if (data1[k,b]<knot2[i,j]) data2[k,j]=0 else data2[k,j]=data1[k,b]-
      knot2[i,j]
    }
  }
  mx=cbind(aa,data1,data2)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p
  A=mx%*%C%*%t(mx)
  A1=(F-A)
  A2=(sum(diag(A1))/p)^2
  GCV[i]=MSE/A2
}
GCV=as.matrix(GCV)
Rsqr=as.matrix(Rsqr)

```



```

cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (knot2)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (Rsqr)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 2 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d://validasi_GCV2.csv")
write.csv(Rsqr, file="d://validasi_Rsqr2.csv")
write.csv(knot2, file="d://validasi_knot2.csv")

```

Lampiran 4. Program Pemilihan Titik Knot Optimal dengan Tiga Titik Knot Menggunakan Software R

```
GCV3=function(para)
{
  data=read.table("d://data.txt",header=FALSE)
  data=as.matrix(data)
  p=length(data[,1])
  q=length(data[1,])
  m=ncol(data)-para-1
  F=matrix(0,nrow=p,ncol=p)
  dataA=data[, (para+2):q]
  diag(F)=1
  nk= length(seq(min(data[,2]),max(data[,2]),length.out=50))
  knot=matrix(ncol=m,nrow=nk)
  for (i in (1:m))
  {
    for (j in (1:nk))
    {
      a=seq(min(dataA[,i]),max(dataA[,i]),length.out=50)
      knot[j,i]=a[j]
    }
  }
  knot=knot[2:(nk-1),]
  a2=nrow(knot)
  z=(a2*(a2-1)*(a2-2)/6)
  knot1=cbind(rep(NA,(z+1)))
  for (i in (1:m))
  {
    knot2=rbind(rep(NA,3))
    for ( j in 1:(a2-2))
    {
      for (k in (j+1):(a2-1))
      {
        for (g in (k+1):a2)
        {
          xx=cbind(knot[j,i],knot[k,i],knot[g,i])
          knot2=rbind(knot2,xx)
        }
      }
    }
  }
}
```

```

}
}
}
knot1=cbind(knot1,knot2)
}
knot1=knot1[2:(z+1),2:(3*m+1)]
aa=rep(1,p)
data1=matrix(ncol=(3*m),nrow=p)
data2=data[, (para+2):q]
a1=length(knot1[,1])
GCV=rep(NA,a1)
Rsqr=rep(NA,a1)
for (i in 1:a1)
{
  for (j in 1:ncol(knot1))
  {
    b=ceiling(j/3)
    for (k in 1:p)
    {
      if (data2[k,b]<knot1[i,j]) data1[k,j]=0 else data1[k,j]=data2[k,b]-
      knot1[i,j]
    }
  }
  mx=cbind(aa,data[,2:q],data1)
  mx=as.matrix(mx)
  C=pinv(t(mx)%*%mx)
  B=C%*%(t(mx)%*%data[,1])
  yhat=mx%*%B
  SSE=0
  SSR=0
  for (r in (1:p))
  {
    sum=(data[r,1]-yhat[r,])^2
    sum1=(yhat[r,]-mean(data[,1]))^2
    SSE=SSE+sum
    SSR=SSR+sum1
  }
  Rsqr[i]=(SSR/(SSE+SSR))*100
  MSE=SSE/p

```

```

cat("=====", "\n")
cat("Nilai Knot dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (knot1)
cat("=====", "\n")
cat("Rsqr dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (Rsqr)
r=max(Rsqr)
print (r)
cat("=====", "\n")
cat("HASIL GCV dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
print (GCV)
s1=min(GCV)
cat("=====", "\n")
cat("HASIL GCV terkecil dengan Spline linear 3 knot", "\n")
cat("=====", "\n")
cat(" GCV =", s1, "\n")
write.csv(GCV, file="d://validasi_GCV3.csv")
write.csv(Rsqr, file="d://validasi_Rsqr3.csv")
write.csv(knot1, file="d://validasi_knot3.csv")
}

```

Lampiran 5.Program Pemilihan Titik Knot Optimal Empat
Varibel dengan Kombinasi Titik Knot
Menggunakan *Software R*

```
GCVkom=function(para)
{
data=read.table("d://data.txt",header=FALSE)
data=as.matrix(data)
p1=length(data[,1])
q1=length(data[1,])
v=para+2
F=matrix(0,nrow=p1,ncol=p1)
diag(F)=1
x1=read.table("d://knotx1.txt")
x2=read.table("d://knotx2.txt")
x3=read.table("d://knotx3.txt")
x4=read.table("d://knotx4.txt")
n2=nrow(x1)
a=matrix(nrow=4,ncol=3^4)
m=0
for (ii in 1:3)
for (j in 1:3)
for (k in 1:3)
for (l in 1:3)
{
m=m+1
a[,m]=c(ii,j,k,l)
}
a=t(a)
GCV=matrix(nrow=nrow(x1),ncol=3^4)
R=matrix(nrow=nrow(x1),ncol=3^4)
for (i in 1:3^4)
{
for (h in 1:nrow(x1))
{
if (a[i,1]==1)
{
gab=as.matrix(x1[,1])
gen=as.matrix(data[,v])
```

```

aa=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,1]==2)
{
gab=as.matrix(x1[,2:3])
gen=as.matrix(cbind(data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1: nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x1[,4:6])
gen=as.matrix(cbind(data[,v],data[,v],data[,v]))
aa=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) aa[w,j]=0 else aa[w,j]=gen[w,j]-gab[h,j]
}
}
}
if (a[i,2]==1)
{
gab=as.matrix(x2[,1] )
gen=as.matrix(data[, (v+1)])
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{

```

```

if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
if (a[i,2]==2)
{
gab=as.matrix(x2[,2:3] )
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x2[,4:6])
gen=as.matrix(cbind(data[, (v+1)],data[, (v+1)],data[, (v+1)]))
bb=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) bb[w,j]=0 else bb[w,j]=gen[w,j]-gab[h,j]
}
}
if (a[i,3]==1)
{
gab=as.matrix(x3[,1] )
gen=as.matrix(data[, (v+2)])
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
}

```

```

else
if (a[i,3]==2)
{
gab=as.matrix(x3[,2:3] )
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
{
gab=as.matrix(x3[,4:6])
gen=as.matrix(cbind(data[, (v+2)],data[, (v+2)],data[, (v+2)]))
cc=matrix(nrow=nrow(x1)*nrow(data),ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) cc[w,j]=0 else cc[w,j]=gen[w,j]-gab[h,j]
}
}
}
if (a[i,4]==1)
{
gab=as.matrix(x4[,1] )
gen=as.matrix(data[, (v+3)])
dd=matrix(nrow=nrow(x1)*nrow(data),ncol=1)
for (j in 1:1)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
}
else
if (a[i,4]==2)
{

```



```

gab=as.matrix(x4[,2:3] )
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=2)
for (j in 1:2)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
else
{
gab=as.matrix(x4[,4:6])
gen=as.matrix(cbind(data[, (v+3)], data[, (v+3)], data[, (v+3)]))
dd=matrix(nrow=nrow(x1)*nrow(data), ncol=3)
for (j in 1:3)
for (w in 1:nrow(data))
{
if (gen[w,j]<gab[h,j]) dd[w,j]=0 else dd[w,j]=gen[w,j]-gab[h,j]
}
}
ma=as.matrix(cbind(aa,bb,cc,dd))
mx=cbind(rep(1,nrow(data)), data[,2:q1], na.omit(ma))
mx=as.matrix(mx)
C=pinv(t(mx)%*%mx)
B=C%*%(t(mx)%*%data[,1])
yhat=mx%*%B
SSE=0
SSR=0
for (r in 1:nrow(data))
{
sum=(data[r,1]-yhat[r,])^2
sum1=(yhat[r,]-mean(data[,1]))^2
SSE=SSE+sum
SSR=SSR+sum1
}
Rsqr=(SSR/(SSE+SSR))*100
R[i]=Rsqr
MSE=SSE/p1

```

```

A=mx%%C%%t(mx)
A1=(F-A)
A2=(sum(diag(A1))/p)^2
GCV[h,i]=MSE/A2
}
if (a[i,1]==1) sp=x1[,1] else
if (a[i,1]==2) sp=x1[,2:3] else
sp=x1[,4:6]
if (a[i,2]==1) spl=x2[,1] else
if (a[i,2]==2) spl=x2[,2:3] else
spl=x2[,4:6]
if (a[i,3]==1) splin=x3[,1] else
if (a[i,3]==2) splin=x3[,2:3] else
splin=x3[,4:6]
if (a[i,4]==1) spline=x4[,1] else
if (a[i,4]==2) spline=x4[,2:3] else
spline=x4[,4:6]
kkk=cbind(sp,spl,splin,spline)
cat("=====","\n")
print(i)
print(kkk)
print(Rsq)
}
write.csv(GCV,file="d://output GCVkom.csv")
}

```

Lampiran 6. Program Estimasi Parameter dengan Kombinasi Titik Knot

```

uji=function(alpha,para)
{
data=read.table("d://data.txt",header=FALSE)
knot=read.table("d://KNOT.txt",header=FALSE)
data=as.matrix(data)
knot=as.matrix(knot)
ybar=mean(data[,1])
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+1],d
ata[,m+2], data[,m+3],data[,m+3],data[,m+3])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-
knot[1,i]
}
}
mx=cbind(satu,data[,2],data.knot[,1],data[,3],data.knot[,2:4],data[,4],d
ata.knot[,5],data[,5],data.knot[,6:8])
mx=as.matrix(mx)
B=(pinv(t(mx)%*%mx))%*%t(mx)%*%data[,1]
cat("=====", "\n")
cat("Estimasi Parameter", "\n")
cat("=====", "\n")
print (B)
n1=nrow(B)
yhat=mx%*%B

```

```

res=data[,1]-yhat
SSE=sum((data[,1]-yhat)^2)
SSR=sum((yhat-ybar)^2)
SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsq=(SSR/(SSR+SSE))*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Tolak Ho yakni minimal terdapat 1 prediktor yang
signifikan", "\n")
cat("", "\n")
}
else
{
cat("-----", "\n")
cat("Kesimpulan hasil uji serentak", "\n")
cat("-----", "\n")
cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan", "\n")
cat("", "\n")
}

#uji t (uji individu)
thit=rep(NA,n1)
pval=rep(NA,n1)
SE=sqrt(diag(MSE*(pinv(t(mx)%*%mx))))
cat("-----", "\n")
cat("Kesimpulan hasil uji individu", "\n")
cat("-----", "\n")

```

```

thit=rep(NA,n1)
pval=rep(NA,n1)
for (i in 1:n1)
{
  thit[i]=B[i,1]/SE[i]
  pval[i]=2*(pt(abs(thit[i]),(p-n1),lower.tail=FALSE))
  if (pval[i]<=alpha) cat("Tolak Ho yakni prediktor signifikan dengan
  pvalue",pval[i],"\n") else cat("Gagal tolak Ho yakni prediktor tidak
  signifikan dengan pvalue",pval[i],"\n")
}
thit=as.matrix(thit)
cat("=====", "\n")
cat("nilai t hitung", "\n")
cat("=====", "\n")
print (thit)
cat("Analysis of Variance", "\n")
cat("=====", "\n")
cat("Sumber      df      SS      MS      Fhit", "\n")
cat("Regresi      ",(n1-1), " ",SSR, " ",MSR, " ",Fhit, "\n")
cat("Error        ",p-n1, " ",SSE, " ",MSE, "\n")
cat("Total        ",p-1, " ",SST, "\n")
cat("=====", "\n")
cat("s=",sqrt(MSE), "      Rsq=",Rsq, "\n")
cat("pvalue(F)=",pvalue, "\n")
write.csv(res,file="d://output uji residual.csv")
write.csv(pval,file="d://output uji pvalue.csv")
write.csv(mx,file="d://output uji mx.csv")
write.csv(yhat,file="d://output uji yhat.csv")
write.csv(B,file="d://output B.csv")
write.csv(thit,file="d://output thit.csv")

```

Lampiran 7. Program Uji Glejser untuk Kombinasi Titik Knot

```

glejser=function(alpha,para)
{
data=read.table("d://VALIDASI_TA//data2.txt")
knot=read.table("d://VALIDASI_TA//KNOT.txt",header=FALSE)
res=read.table("d://VALIDASI_TA//residual.txt")
data=as.matrix(data)
knot=as.matrix(knot)
res=abs(res)
res=as.matrix(res)
rbar=mean(res)
m=para+2
p=nrow(data)
q=ncol(data)
dataA=cbind(data[,m],data[,m+1],data[,m+1],data[,m+1],data[,m+1],
data[,m+2], data[,m+3],data[,m+3],data[,m+3])
dataA=as.matrix(dataA)
satu=rep(1,p)
n1=ncol(knot)
data.knot=matrix(ncol=n1,nrow=p)
for (i in 1:n1)
{
for(j in 1:p)
{
if (dataA[j,i]<knot[1,i]) data.knot[j,i]=0 else data.knot[j,i]=dataA[j,i]-
knot[1,i]
}
}
mx=cbind(satu,
data[,2],data.knot[,1],data[,3],data.knot[,2:4],data[,4],data.knot[,
5],data[,5],data.knot[,6:8])
mx=as.matrix(mx)
B=(ginv(t(mx)%*%mx))%*%t(mx)%*%res
n1=nrow(B)
yhat=mx%*%B
SSE=sum((res-yhat)^2)
SSR=sum((yhat-rbar)^2)

```

```

SST=SSR+SSE
MSE=SSE/(p-n1)
MSR=SSR/(n1-1)
Rsqr=(SSR/SST)*100

#uji F (uji serentak)
Fhit=MSR/MSE
pvalue=pf(Fhit,(n1-1),(p-n1),lower.tail=FALSE)
if (pvalue<=alpha)
{
  cat("-----","\n")
  cat("Kesimpulan hasil uji serentak","\n")
  cat("-----","\n")
  cat("Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan
atau terjadi heteroskedastisitas","\n")
  cat("", "\n")
}
else
{
  cat("-----","\n")
  cat("Kesimpulan hasil uji serentak","\n")
  cat("-----","\n")
  cat("Gagal Tolak Ho yakni semua prediktor tidak berpengaruh
signifikan atau tidak terjadi heteroskedastisitas","\n")
  cat("", "\n")
}
cat("Analysis of Variance", "\n")
cat("=====","\n")
cat("Sumber    df    SS    MS    Fhit","\n")
cat("Regresi    ",(n1-1)," ",SSR," ",MSR,"",Fhit,"\n")
cat("Error      ",p-n1," ",SSE,"",MSE,"\n")
cat("Total      ",p-1," ",SST,"\n")
cat("=====","\n")
cat("s=",sqrt(MSE),"    Rsqr=",Rsqr,"\n")
cat("pvalue(F)=",pvalue,"\n")
}

```

Lampiran 8. Output Nilai GCV dengan Satu Titik Knot Pada Empat Variabel Prediktor

No	X ₁	X ₂	X ₃	X ₄	GCV
1	11,8	6,46	0,33	7,93	4,44
2	12,3	6,57	0,34	8,02	4,34
3	12,81	6,69	0,35	8,11	4,22
4	13,31	6,8	0,35	8,2	4,1
5	13,81	6,91	0,36	8,29	4,07
6	14,31	7,03	0,37	8,38	4,1
7	14,82	7,14	0,38	8,47	4,14
8	15,32	7,25	0,39	8,56	4,18
9	15,82	7,37	0,39	8,65	4,26
10	16,33	7,48	0,4	8,75	4,37
⋮	⋮	⋮	⋮	⋮	⋮
48	27,90	10,07	0,59	10,83	5,79

Lampiran 9.Output Nilai GCV dengan Dua Titik Knot Pada Empat Variabel Prediktor

No	X ₁	X ₂	X ₃	X ₄	GCV
1	3,75	4,66	0,20	6,48	4,44
	11,80	6,46	0,33	7,93	
2	3,75	4,66	0,20	6,48	4,22
	12,81	6,69	0,35	8,11	
3	3,75	4,66	0,20	6,48	4,10
	13,31	6,80	0,35	8,20	
4	3,75	4,66	0,20	6,48	4,07
	13,81	6,91	0,36	8,29	
5	3,75	4,66	0,20	6,48	4,10
	14,31	7,03	0,37	8,38	
⋮	⋮	⋮	⋮	⋮	⋮
1225	3,75	4,66	0,20	6,48	4,49
	16,83	7,59	0,41	8,80	

Lampiran 10.Output Nilai GCV dengan Tiga Titik Knot Pada Empat Variabel Prediktor

No	X ₁	X ₂	X ₃	X ₄	GCV
1	17,33	7,70	0,42	8,93	3,85
	18,84	8,04	0,44	9,20	
	20,35	8,38	0,47	9,47	
2	17,33	7,70	0,42	8,93	2,88
	18,84	8,04	0,44	9,20	
	20,85	8,49	0,48	9,56	
3	17,33	7,70	0,42	8,93	2,63
	18,84	8,04	0,44	9,20	
	21,36	8,61	0,49	9,65	
4	17,33	7,70	0,42	8,93	2,97
	18,84	8,04	0,44	9,20	
	21,86	8,72	0,49	9,74	
5	17,33	7,70	0,42	8,93	3,52
	18,84	8,04	0,44	9,20	
	22,36	8,83	0,50	9,83	
⋮	⋮	⋮	⋮	⋮	⋮
17296	17,33	7,70	0,42	8,93	3,77
	18,84	8,04	0,44	9,20	
	23,87	9,17	0,53	10,10	

Lampiran 11,Output Nilai GCV dengan Kombinasi Titik Knot
Pada Empat Variabel Prediktor

No	X ₁	X ₂	X ₃	X ₄	GCV
1	13,81	4,66	0,42	6,48	3,53
		6,91	0,44	8,29	
			0,49		
2	13,81	4,66	0,42	8,93	2,72
		6,91	0,44	9,2	
			0,49	9,65	
3	13,81	7,7	0,36	8,29	4,57
		8,04			
		8,61			
4	13,81	7,7	0,36	8,93	2,36
		8,04		9,2	
		8,61		9,65	
5	13,81	7,70	0,36	6,48	4,57
		8,04		8,29	
		8,61			
⋮	⋮	⋮	⋮	⋮	⋮
81	3,75	6,91	0,36	8,29	4,07
	13,81				

Lampiran 12. Output Nilai GCV dengan Satu Titik Knot Pada Dua Variabel Prediktor

No	X ₂	X ₄	GCV
1	6,46	7,93	4,09
2	6,57	8,02	3,93
3	6,69	8,11	3,78
4	6,8	8,2	3,62
5	6,91	8,29	3,56
6	7,03	8,38	3,57
7	7,14	8,47	3,6
8	7,25	8,56	3,64
9	7,37	8,65	3,7
10	7,48	8,75	3,79
⋮	⋮	⋮	⋮
48	10,07	10,83	5,24

Lampiran 13.Output Nilai GCV dengan Dua Titik Knot Pada
Dua Variabel Prediktor

No	X ₂	X ₄	GCV
1	8,27	9,38	4,33
	9,85	10,65	
2	8,27	9,38	4,33
	9,96	10,74	
3	8,27	9,38	4,33
	10,07	10,83	
4	8,27	9,38	4,45
	10,19	10,92	
5	8,38	9,47	3,66
	8,49	9,56	
⋮	⋮	⋮	⋮
1225	8,38	9,47	4,42
	9,06	10,01	

Lampiran 14, Output Nilai GCV dengan Tiga Titik Knot Pada Dua Variabel Prediktor

No	X ₂	X ₄	GCV
1	7,7	8,93	3,31
	8,04	9,2	
	10,07	10,83	
2	7,7	8,93	2,69
	8,16	9,29	
	8,27	9,38	
3	7,7	8,93	2,47
	8,16	9,29	
	8,38	9,47	
4	7,7	8,93	1,95
	8,16	9,29	
	8,49	9,56	
5	7,7	8,93	1,78
	8,16	9,29	
	8,61	9,65	
⋮	⋮	⋮	⋮
17296	7,7	8,93	3,26
	8,16	9,29	
	8,95	9,92	

Lampiran 15. Output Nilai GCV dengan Kombinasi Titik Knot
Pada Dua Variabel Prediktor

No	X ₂	X ₄	GCV
1	8,38	8,29	4,76
	8,61		
2	8,38	9,47	3,51
	8,61	9,65	
3	8,38	8,93	2,38
	8,61	9,29	
		9,65	
4	7,7	8,29	3,77
	8,16		
	8,61		
5	7,7	9,47	2,47
	8,16	9,65	
	8,61		
⋮	⋮	⋮	⋮
9	7,7	8,93	1,78
	8,16	9,29	
	8,61	9,65	

Lampiran 16. Output Estimasi Parameter dan Uji Signifikansi Parameter Model Pada Empat Variabel Prediktor

=====

Estimasi Parameter

=====

[,1]
 [1,] -1.709691768
 [2,] 0.006715846
 [3,] -0.014966207
 [4,] 1.425813271
 [5,] -8.011831812
 [6,] 11.872440942
 [7,] -10.318110212
 [8,] 7.894642444
 [9,] -13.588194906
 [10,] -0.458890769
 [11,] 26.218264824
 [12,] -48.650180578
 [13,] 24.941693102

Kesimpulan hasil uji serentak

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

Kesimpulan hasil uji individu

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.7486949
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.9365259
 Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.9153273
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.002386174
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.0204703
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.02268266
 Tolak Ho yakni prediktor signifikan dengan pvalue 0.009783395

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.201954

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.2298841

Gagal tolak H_0 yakni prediktor tidak signifikan dengan pvalue 0.4176313

Tolak H_0 yakni prediktor signifikan dengan pvalue 0.0003629487

Tolak H_0 yakni prediktor signifikan dengan pvalue 7.123153e-05

Tolak H_0 yakni prediktor signifikan dengan pvalue 4.573866e-05

=====

nilai t hitung

=====

[,1]

[1,] -0.3246002

[2,] 0.0805966

[3,] -0.1076091

[4,] 3.4522619

[5,] -2.5069082

[6,] 2.4592907

[7,] -2.8410898

[8,] 1.3172305

[9,] -1.2366098

[10,] -0.8268273

[11,] 4.2435236

[12,] -4.9269693

[13,] 5.1147018

Analysis of Variance

=====

Sumber	df	SS	MS	Fhit
Regresi	12	123.5773	10.29811	7.060777
Error	21	30.6284	1.458495	
Total	33	154.2057		

=====

s= 1.207682 Rsq= 80.13796

pvalue(F)= 5.713139e-05 pvalue(F)= 0.0001849592

Lampiran 17. Output Estimasi Parameter dan Uji Signifikansi Parameter Model Pada Dua Variabel Prediktor

Estimasi Parameter

```

=====
[.,1]
[1,] -1.4808664
[2,]  1.7071706
[3,] -7.4142017
[4,] 11.9968100
[5,] -11.3672083
[6,] -0.4087986
[7,] 15.6019816
[8,] -39.6194422
[9,] 26.5727251
=====

```

Kesimpulan hasil uji serentak

Tolak Ho yakni minimal terdapat 1 prediktor yang signifikan

Kesimpulan hasil uji individu

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.7250564

Tolak Ho yakni prediktor signifikan dengan pvalue 3.520487e-05

Tolak Ho yakni prediktor signifikan dengan pvalue 0.002802836

Tolak Ho yakni prediktor signifikan dengan pvalue 0.004281853

Tolak Ho yakni prediktor signifikan dengan pvalue 0.0007145665

Gagal tolak Ho yakni prediktor tidak signifikan dengan pvalue 0.3504906

Tolak Ho yakni prediktor signifikan dengan pvalue 0.0001672715

Tolak Ho yakni prediktor signifikan dengan pvalue 6.302501e-06

Tolak Ho yakni prediktor signifikan dengan pvalue 3.295731e-06

=====

nilai t hitung

=====

[,1]

[1,] -0.3556902

[2,] 5.0220472

[3,] -3.3145073

[4,] 3.1420059

[5,] -3.8570513

[6,] -0.9514382

[7,] 4.4206674

[8,] -5.6916184

[9,] 5.9472705

Analysis of Variance

=====

Sumber	df	SS	MS	Fhit
Regresi	8	121.4208	15.1776	11.57363
Error	25	32.78488	1.311395	
Total	33	154.2057		

=====

s= 1.145162 Rsq= 78.73952

pvalue(F)= 1.031246e-06

Lampiran 18. Output Uji Glejser

Kesimpulan hasil uji serentak

Gagal Tolak H_0 yakni semua prediktor tidak berpengaruh signifikan atau tidak terjadi heteroskedastisitas

Analysis of Variance

=====

Sumber	df	SS	MS	Fhit
Regresi	8	2.57228	0.321535	0.8606192
Error	25	9.340222	0.3736089	
Total	33	11.9125		

=====

$s = 0.6112356$ $Rsq = 21.59311$

$pvalue(F) = 0.5611292$

(Halaman ini sengaja dikosongkan)

Lampiran 19. Surat Pernyataan Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMIPA ITS :

Nama : Hanida Muryaninggar

NRP : 1313100090

Menyatakan bahwa data yang digunakan dalam Tugas Akhir/~~Thesis~~ ini merupakan data sekunder yang diambil dari ~~penelitian/ buku/ Tugas Akhir/ Thesis/~~ publikasi lainnya yaitu :

Sumber : Badan Pusat Statistik

Keterangan : www.bps.go.id

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir

Surabaya, Juli 2017

(Prof. Dr. I Nyoman Budiantara., M.Si)
NIP. 19650603 198903 1 003

(Hanida Muryaninggar)
NRP. 1313100090

BIODATA PENULIS



Penulis dengan nama lengkap Hanida Muryaninggar dan nama panggilan Hanida dengan tempat tanggal lahir Surabaya, 10 Juli 1995. Bungsu dari dua bersaudara pasangan Bapak Toha dan Ibu Ida. Penulis memiliki hobi travelling mendengarkan music, menonton film, dan membaca. Penulis menempuh pendidikan formal di SDN Keputih 245 Surabaya tahun 2001-2007, SMP Muhammadiyah 5 Surabaya tahun 2007-2010, SMA Muhammadiyah 2 Surabaya tahun

2010-2013. Pada tahun 2013 penulis diterima menjadi mahasiswa Jurusan Statistika di Institut Teknologi Sepuluh Nopember (ITS) melalui jalur SBMPTN. Selama menempuh pendidikan S1, penulis pernah mendapatkan beasiswa PPA (Peningkatan Prestasi Akademik) pada tahun ke tiga. Kegiatan kepanitiaan yang pernah diikuti baik kepanitiaan institut dan jurusan yaitu Bursa Karir ITS dari tahun 2014-2017 dan Pekan Raya Statistik tahun 2015. Untuk berdiskusi lebih lanjut mengenai tugas akhir, hubungi penulis melalui:

Email: hanidha.murya@gmail.com